ience
brary

173

JUL 6 1937

Vol. 21, No. 1 July, 1937

PSYCHOLOGICAL REVIEW PUBLICATIONS

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER

UNIVERSITY OF PENNSYLVANIA

CONTENTS

The Effects of Oxygen Deprivation on Eye Movements in Reading: R. A. McFARLAND, C. A. KNEHR AND C. BERENS 1

The Effect of Induced Muscular Tension Upon Various Phases of the Learning Process: J. C. STAUFFACHER 26

The Generalization of Conditioned Responses. III. Extinction, Spontaneous Recovery, and Disinhibition of Conditioned and of Generalized Responses: C. I. HOVLAND 47

An Ontogenetic Study of Cerebral Electrical Potentials in the Guinea Pig: H. H. JASPER, C. S. BRIDGMAN AND L. CARMICHAEL 63

Cardiac Acceleration in Emotional Situations: J. G. BEEBE-CENTER AND S. S. STEVENS 72

Studies on the Physiology of Sleep. Changes in Irritability to Auditory Stimuli During Sleep: F. J. MULLIN, N. KLEITMAN AND N. R. COOPERMAN 88

Blocking in Mental and Motor Tasks During a 65-Hour Vigil: N. WARREN AND B. CLARK 97

A Comparison of Collective and Individual Judgments of Fact: H. GURNEE .. 106

Confidence Rating as an Index of Difficulty: M. HERTZMAN 113

Report of a Minor Investigation of Extra-Sensory Perception: K. H. BAKER 120

PUBLISHED MONTHLY, TWO VOLUMES PER YEAR
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, University of Pennsylvania (*J. Exper. Psych.*)
WALTER S. HUNTER, Brown University (*Abstracts*)
HENRY T. MOORE, Skidmore College (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, Princeton University (*Review*)
JOHN A. McGEOCH, Wesleyan University (*Bulletin*)
JOHN F. DASHIELL, Univ. of North Carolina (*Monographs*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the Bulletin consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears monthly (beginning January, 1937), two volumes per year, each volume of six numbers containing about 625 pages.

PSYCHOLOGICAL INDEX

(discontinued after No. 42, June, 1936)

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The Index has been published annually from 1895 to 1936, the volumes covering the literature of 1894 through 1935. Its bibliographical services will be carried on by the Psychological Abstracts.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The Monographs appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, June, September, December, March, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Journal:** $14.00 ($7.00 per volume, Foreign $7.25).
Monographs: $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $7.00 (Foreign, $7.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: **Journal,** $1.25; **Review,** $1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp. (2 vols.): $17.00 (Foreign, $17.75).
Bulletin and J. Exp. (2 vols.): $18.00 (Foreign, $18.75).
Review, Bulletin, and J. Exp. (2 vols.): $22.00 (Foreign, $23.00).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. 21, NO. 1 JULY, 1937

THE EFFECTS OF OXYGEN DEPRIVATION ON EYE MOVEMENTS IN READING *

BY ROSS A. McFARLAND, PH.D., CHAS. A. KNEHR, M.A.
AND CONRAD BERENS, M.D.

INTRODUCTION

A marked variation in any of the organic constants such as the temperature, the electrolytes or the oxygen saturation produces striking effects on the central nervous system. The impairment due to oxygen lack has been frequently observed on mountain expeditions, during balloon and aeroplane ascents, and in simulated high altitude at sea level (chambers and rebreathing devices) (*cf.* 11). During prolonged exposures the effects are partially compensated for by acclimatization (*i.e.*, changes in the blood, the circulation, and the respiration) and during short experiments the impairment is often concealed by exerting greater effort or in the practice effects involved in most tests. Since the eye movements in reading are a highly practiced type of motor activity showing little variability under normal conditions it was believed that the measurement of these reactions might serve as a sensitive test of the effects of oxygen lack on the nervous system. The object of this study, therefore, was to analyze not only the effects of oxygen want on an unconscious and well established motor habit but also the basic mechanisms involved in read-

* From the Department of Psychology, Columbia University, and the New York Eye and Ear Infirmary. The authors wish to express their indebtedness to the Linde Air Products Company for the gases used in the research and to the Ophthalmological Foundation, Incorporated, for the use of apparatus and financial assistance in publication.

ing. By placing the organism under stress (following a complete eye examination) it was believed that known defects might be accentuated, latent anomalies revealed, and certain basic ocular functions more precisely measured.

The actual character of the reading movements was probably first observed by Javal (**10**) in 1879. He noted that the eyes did not move uninterruptedly over the printed line but made a succession of pauses. From the early investigations of Erdmann, Dodge, Huey, and Dearborn came the method of photographing the points of light reflected from the eccentric curvature of the corneas on a moving film while the eyes were reading. The more recent techniques are essentially the same as those developed by the early investigators. Certain refinements have been made, however, by Buswell, Judd, Gray, Schmidt, and others, making possible the study of ocular movements in two dimensions, with more exact timing, improved photographic technique and continuous recording. All phases of this problem have been discussed in detail by Tinker (**19**), Taylor (**16**), and Vernon (**20**).

Dodge and Benedict (**7**) studied the effects of alcohol on the ocular movements involved in fixating, as rapidly as possible, typewritten letters exposed in any one of 6 positions. They found that 30 cc of alcohol seemed to decrease the latency of the eye movements, while 45 cc increased it. When fixating two letters alternately, as fast as possible, there was a decrease in speed of 2.5 percent with 30 cc of alcohol and 18 percent following the ingestion of 45 cc. Scofield (**17**) has recently reported a study of the effects of small amounts of alcohol and caffeine on optic nystagmus (Dodge mirror-recording technique). The pursuit movements were shortened and in some cases they disappeared, especially with the larger amounts of alcohol. The organism apparently more readily abandons the attempt to adjust when alcohol has been ingested than under normal conditions.

The use of reduced oxygen concentrations to bring about alterations in the internal environment is more direct and basic than alcohol in that nervous functioning depends upon an adequate oxygen supply. McFarland and Barach (**13**),

and others, have shown that the effects of alcohol are somewhat similar to those of oxygen deprivation since alcohol suppresses oxidation, especially in the nervous tissues. The oxygen reduction method would tend to minimize the psychological effects produced since the oxygen concentrations can easily be changed unknown to the subject. In testing pilots for altitude fitness during the Great War in the rebreather and low pressure chamber, Wilmer and Berens (**22**) revealed an impairment of the judgment of distance, size of visual fields, accommodation, convergence, and retinal sensitivity with suddenly induced anoxemia. Gellhorn and Spiesman (**8**) report a lengthening of the latent period for negative after-images, a loss of visual intensity and discrimination, and the presence of nystagmus 10 to 30 minutes following inhalation of 9–10 percent oxygen (19–22,000 ft). Similar findings are reported by Goldmann and Schubert (**9**). During gradual adaptation to altitudes as high as 20,000 feet in the Andes, McFarland (**12**) has shown that there is a variation in the phorias, an increased fatiguability of the muscles of accommodation and convergence, and a lengthening of the latent period of negative after-images. The psychological effects of oxygen deprivation have been reviewed in detail by McFarland (**11, 12**).

Experimental Procedure

This experiment was carried out in a Barach portable oxygen chamber (measuring $7 \times 8 \times 8$ ft) in which the partial pressure of gases as well as the temperature and the humidity can be controlled. The partial pressure or percentage of oxygen in the chamber is reduced by diluting the air with nitrogen. The control of the humidity and temperature is maintained by a motor blower unit which contains ice to cool and dry the air. While ventilating the chamber, the relative humidity was kept between 35 and 45 percent and the temperature between 65° and 75° F. The chamber air was forced through a pipe on the blower intake containing soda lime to absorb the excess CO_2. The percentages of O_2 and CO_2 were checked frequently. The CO_2 never exceeded 0.7 percent. The experimenter received an adequate oxygen supply by means of a catheter delivering 2–3 liters of O_2 per minute.

The ten subjects used in this experiment were college students 19–34 years of age. Each subject was tested three times during the experiment: (*a*) under control conditions (air), (*b*) 12.5 percent O_2 and (*c*) 10.5 percent O_2 corresponding to altitudes of sea level, 13,500 feet and 18,000 feet respectively (Bureau of Standards data—standard atmosphere corrected to constant temperature of 15° C.). Four of the subjects were also tested in 5 percent CO_2.

A complete eye examination was given to each subject since variations in the ocular movements in reading are related to the refractive and muscular conditions of the eyes. Most experimenters have had difficulty in interpreting the data from photographic records of eye movements in reading due to their failure to take disturbances of ocular function and errors of refraction into consideration previous to the experiment.

The apparatus used for photographing the eye movements is the Opthalmograph.* The corneal reflections were photographed on 35 mm non-halation motion picture film. The camera has the advantage of being easy to operate, is simplified in design, and portable.

The machine is so arranged that the distance from the eye to the reading matter is 33 cm (our measurements). The length of the printed line was 9.7 cm. An angular sweep of 16° 44′ was made by the eyes from the beginning to the end of the line. This angular movement corresponds roughly to a linear displacement on the film of 7 mm, *i.e.*, each millimeter on the film corresponds to an angular movement of the visual axes of more than 2 degrees.

The lights, for producing both illumination of reading material and corneal reflections, are located to the side and approximately in the plane of the horizontal movements for reading, both being below the axis of the recording system. The light rays reaching the eyes are not parallel. The corneal reflections, which are photographed, are projected (upon fixation of the center of reading card) temporally and slightly above the optical axes on the corneas of the two eyes. When the eyes move to successively lower positions, while reading, these points of reflection creep around upward and nasally on the cornea, coming closer together, causing the actual photographic record to widen. This was checked by having a subject read a series of numbers from above downward on a card. There was a widening of .5 to .6 mm between the two tracks corresponding to more than a degree of relative movement of the eyes. Convergence is indicated but not really present and thus an artifact is introduced into the measurements. This was taken into account in interpreting the data. Any changes in the eye record while the eye movements are in the horizontal plane are probably free from artifacts, or, if present, would tend to be symmetrically disposed; and our results indicate no apparent symmetry of data from the beginning to the end of a single line.

The general procedure for testing each subject was as follows: The experimental period was one hour in length. During this hour, the subject read six paragraphs while his eyes were photographed. The first two were read 15 minutes after entering the chamber, the second pair after one half hour, and the third at the end of an hour. This same procedure was maintained throughout the experiment. The subjects were not allowed to read between the tests while in the chamber.

The reading material consisted of a series of 18 paragraphs, eight lines in length. The first and last lines were not included in the measurements. The lines consisted of 53 letter spaces with a range of from 52 to 62 words per six lines measured. All paragraphs were taken from the same work by the same author. The paragraphs selected were of general interest, written in the same style, covered about the same number of points, and lacked technical material. To equate these paragraphs in any more substantial way would not assure one that they would remain equated for this particular group of subjects. Since six paragraphs were read under each experimental condition, slight differences in difficulty would tend to be canceled.

* Manufactured by the American Optical Company (2, 16).

The subject was told that after reading each paragraph he would be required to answer three questions regarding the content of the paragraph. The comprehension was then rated as good, medium, or poor. It was easy to detect whether the comprehension was good or poor, but the intermediate points were at times difficult to judge, so that only one intermediate rating was used.

By means of test-retest methods Tinker (18) has found that 20 or more lines are necessary for individual diagnosis; however, where group comparisons are concerned adequate reliability may be had from reading 5 or 6 lines. He also states that when the material read is strictly comparable, the validity of fixation frequency and perception time is very high and that reading before a camera yields typical performance. In this study 48 lines of comparable material were read under each experimental condition, of which 36 were measured.

The eye movement photographs were enlarged by projection upon a screen which was ruled so that the readings would correspond directly to millimeters on the film. The enlargement was 20 diameters so that a distance of 2 cm. on the screen was labeled 1 mm. In projecting the records, one eye track was centered on the O line of the screen, and then the distance to the other track was measured directly. The error of measurement involved in this procedure was found to be less than 1 percent. More than 20,000 measurements were made from the ocular photographs in this experiment.

The data from the eye movement records and the methods of measuring are explained below and illustrated diagrammatically on a section of M.K.'s record (Fig. 1).

1. *The Average Time for Reading 6 Lines of Print (1 inch or 2.54 cm. of record equals 2 seconds).*—The measurement was made from the beginning of line 2 to the end of line 7 and includes the saccadic movements which occupy roughly 7 percent of the total time (cf. Tinker, 18). The distance (*xy*) would correspond to the time in seconds for reading a single line.

2. *The Average Number of Fixations per Line.*—In the figure there are six fixations in the first line, and six in the second, including the regression.

3. *The Average Number of Regressions.*—A regression is shown in line 2 (Fig. 1).

4. *Line Deviation.*—This is the difference between the position of the eyes relative to one another from the beginning to the end of the line. In the figure this is the distance (*A*) between the tracks compared with the distance (*D*). Whenever *D* is less than *A* the lines on the record approximate, which corresponds to an actual divergence of the optic axes.

5. *Average Binocular Adjustment During the First Fixation.*—This is the binocular relationship for the first fixation expressed by the difference between the distances *A* and *B*. A convergence on the record would make *B* smaller than *A*, and would represent actual divergence of the optic axes.

6. *Average Binocular Adjustment for All Fixations.*—See (5) above.

7. *Binocular Relations During Saccadic Movements.*—Measurements *B* and *C* in the figure.

8. *Qualitative Aspects of the Record.*—Occurrences of specific patterns, nystagmoid movements, and other irregularities.

9. *Rating of the Degree of Comprehension of Material Read.*

10. *Pulse and Blood Pressure.*

There were other measurements which might have been made but the irregularity of the records under anoxemia made this impossible. One relates to a measurement of the angle of the eye tracks during fixations with respect to the direction the film was moving (time axis) in order to determine quantitatively which eye did

more of the adjusting. In Fig. 1 this would correspond to measuring angles P and P'. The adjusting eye would make the larger angle.

Differences were found between the length of the right and left eye tracks along the time axis. In some cases the right eye track for one line was measurably shorter than for the left and vice versa. It is difficult to make accurate vertical measurements, however, on a photographic record designed primarily to measure the horizontal movements.

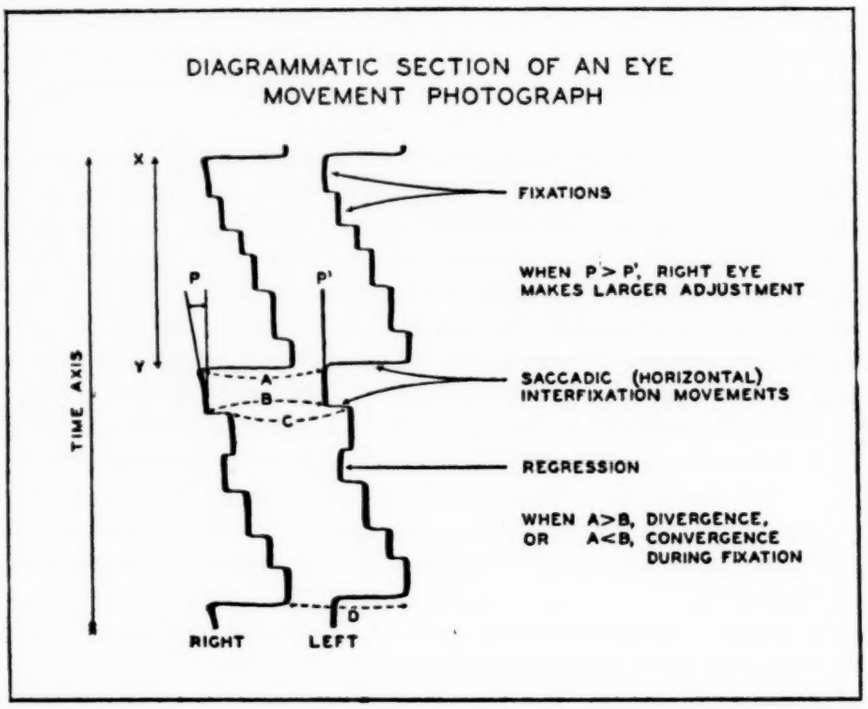

FIG. 1.

RESULTS

The results of this study are presented in Tables 1–8 and Figs. 2 and 3. It was found that the group means mask the effects on the individuals in varying amounts for the different measures. In some cases the group mean follows the individual ones closely and in others the mean remains the same since the individual variations are cancelled.

Time.—The group as a whole showed a significant increase in the average time for reading six lines of print both in 12.5 and 10.5 percent oxygen with an increase in the variability for the group (Table 1 and Fig. 2*A*). The subject who was the fastest reader of the group under control conditions tended to keep his relative position under diminished oxygen, as shown by rank order correlations of .92 and .94.

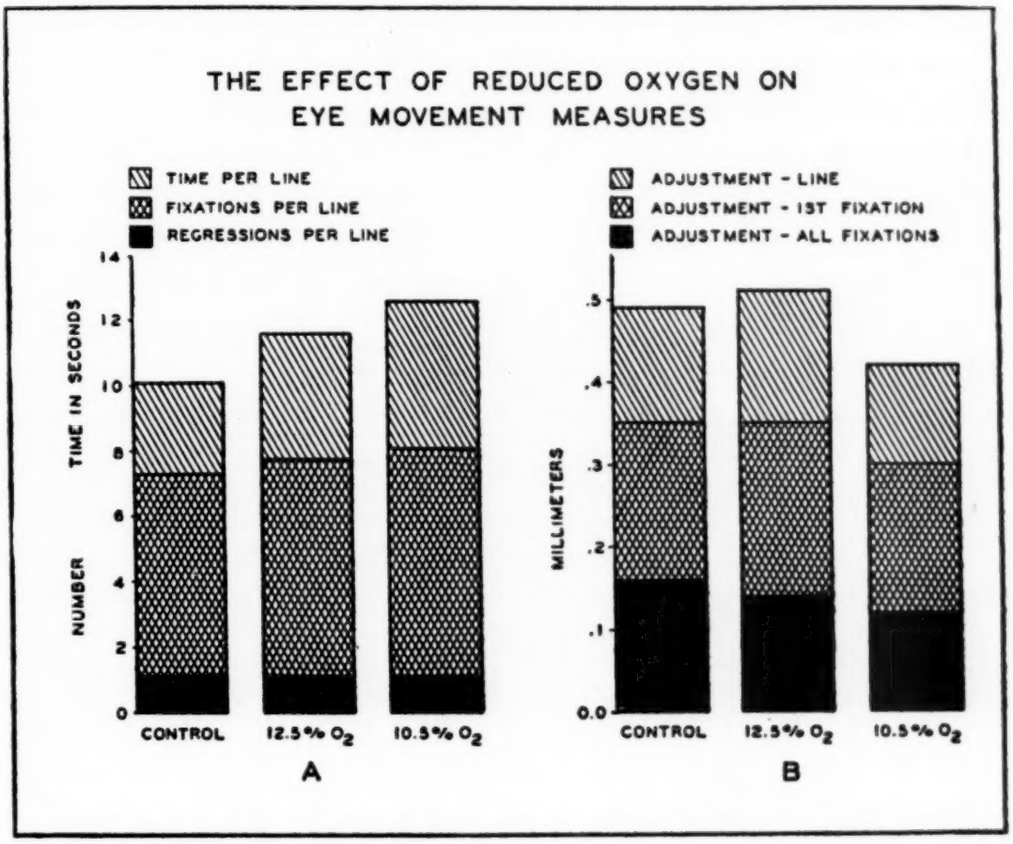
THE EFFECT OF REDUCED OXYGEN ON
EYE MOVEMENT MEASURES
TIME PER LINE
FIXATIONS PER LINE
REGRESSIONS PER LINE
NUMBER TIME IN SECONDS
0
2
4
6
8
10
12
14
CONTROL
12.5% O2
10.5% O2
A
ADJUSTMENT - LINE
ADJUSTMENT - 1ST FIXATION
ADJUSTMENT - ALL FIXATIONS
MILLIMETERS
0.0
.1
.2
.3
.4
.5
CONTROL
12.5% O2
10.5% O2
B

FIG. 2.

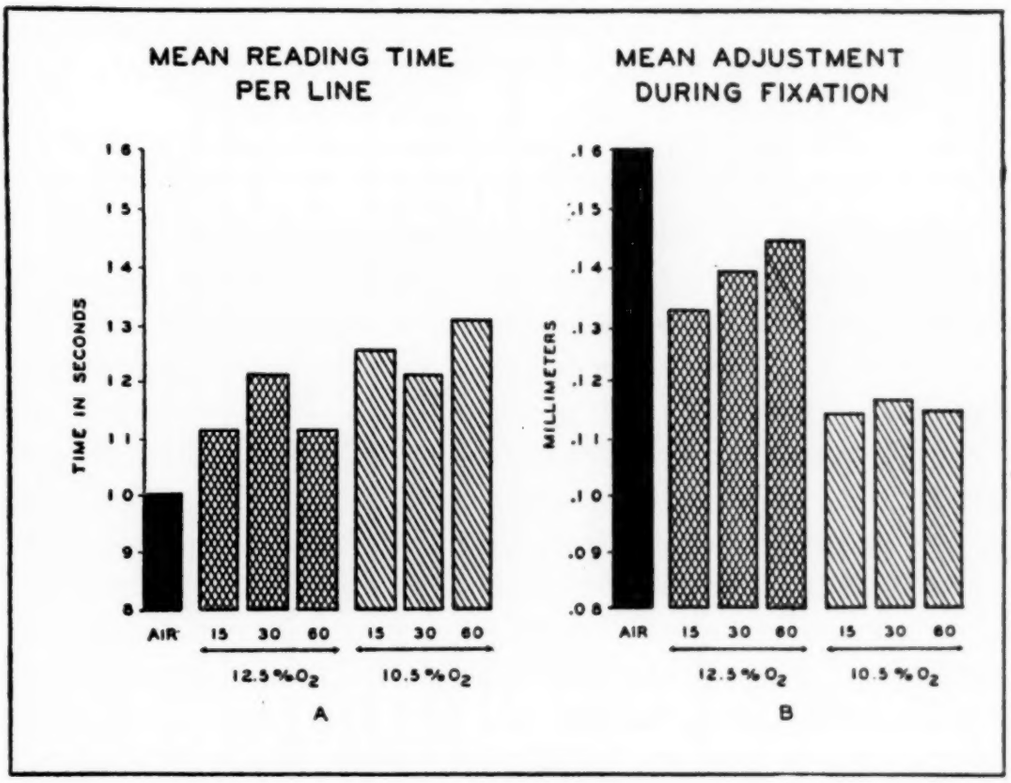
MEAN READING TIME
PER LINE
TIME IN SECONDS
8
9
10
11
12
13
14
15
16
AIR
15
30
60
15
30
60
12.5% O2
10.5% O2
A
MEAN ADJUSTMENT
DURING FIXATION
MILLIMETERS
.08
.09
.10
.11
.12
.13
.14
.15
.16
AIR
15
30
60
15
30
60
12.5% O2
10.5% O2
B

FIG. 3.

TABLE I

THE MEANS, STANDARD DEVIATIONS, CORRELATIONS, AND SIGNIFICANCES OF THE OBSERVED DIFFERENCES FOR THE GROUP FOR ALL VARIABLES IN AIR, 12.5% O_2 AND 10.5% O_2. MEANS AND STANDARD DEVIATIONS ARE SHOWN FOR A GROUP OF FOUR SUBJECTS IN AIR AND 5% CO_2

Variable	Mean			S.D.			Rank Correlations		$D/\sigma D$*		CO_2			
											Mean		S.D.	
	Control	12.5%	10.5%	Control	12.5%	10.5%	Control & 12.5%	Control & 10.5%	Control & 12.5%	Control & 10.5%	Control	5% CO_2	Control	5% CO_2
Mean time for six lines (seconds)...	10.06	11.62	12.61	2.79	3.98	3.95	0.92	0.94	2.65	4.68	10.44	9.95	0.87	0.79
Mean number fixations per line....	7.35	7.76	8.08	1.93	2.07	2.06	0.92	0.95	1.44	3.49	7.89	7.65	1.64	1.32
Mean number regressions per line...	1.10	1.11	1.09	0.62	0.68	0.31	0.96	0.89	—	—	0.99	0.99	0.92	0.76
Mean line divergence (millimeters)..	0.49	0.51	0.42	0.21	0.27	0.31	0.71	0.58	0.34	0.84	0.59	0.49	0.14	0.22
Mean divergence during first fixation (millimeters).....	0.35	0.35	0.30	0.12	0.19	0.19	0.19	0.65	—	0.98	0.44	0.31	0.11	0.13
Mean divergence all fixations (millimeters).....	0.16	0.14	0.12	0.06	0.05	0.05	0.96	0.89	3.02	4.64	0.15	0.12	0.04	0.03

* 2.26 Significant. 3.25 Very significant. (From Fisher.)

If each subject is considered individually (Table 2) the average time for reading six lines increased (with the exception of two subjects) in 12.5 percent O_2 and (with one exception) in 10.5 percent O_2. This increase was statistically reliable for four of the ten subjects in 12.5 percent O_2 and for six of the subjects in 10.5 percent O_2.

Fixations.—The increase for the group in the number of fixations is slight, but significant at 10.5 percent O_2 (Table 1 and Fig. 2*A*). The variability increases only slightly. The various individuals maintain their ranks in the group as shown by correlations of .92 and .95.

TABLE II

AVERAGE TIME IN SECONDS FOR READING SIX LINES

Subject	Control		12.5% O_2		10.5% O_2		$d/\sigma d$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	10.33	1.12	11.03	1.22	11.96	1.03	0.95	2.39
C. K.....	6.53	0.23	6.58	0.42	7.83	0.41	0.21	6.11
R. M.....	10.56	1.00	12.03	0.93	13.84	0.87	2.26	2.91
H. B.....	14.33	1.12	14.64	1.17	15.39	1.20	0.39	1.45
C. S.....	8.44	0.77	11.59	0.52	12.84	1.06	7.59	7.50
W. G.....	12.46	0.65	14.88	1.70	16.50	2.23	2.98	3.91
H. V.....	9.69	0.88	8.88	0.67	9.41	0.61	1.61	0.58
L. R.....	14.44	1.79	20.31	1.95	20.85	1.96	4.96	5.40
W. T.....	6.85	0.87	6.44	0.66	8.61	1.76	0.43	2.05
W. E.....	6.93	1.19	9.80	3.00	8.88	2.05	2.07	1.84

Seven individuals showed an increase in the number of fixations, the others showing little change (Table 3). Three of these increases are statistically significant in 12.5 percent and five are significant in 10.5 percent oxygen.

Regressions.—It was surprising that the average number of regressions per line showed no change for the group (Table 1, Fig. 2*A*). The variability decreased slightly at 10.5 percent. The implications of this will be discussed later. The individual who tends to read with few regressions, as observed in regard to time and number of fixations, maintains his relative position in the group at 12.5 percent O_2 and 10.5 percent O_2 compared with the control series (correlation of .96 and .89 respectively).

TABLE III

MEAN NUMBER OF FIXATIONS PER LINE

Subject	Control		12.5% O_2		10.5% O_2		$d/\sigma d$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	6.89	1.78	7.10	1.47	7.17	1.66	0.55	0.68
C. K.....	5.64	0.63	5.86	0.95	6.39	0.68	1.17	4.87
R. M.....	8.47	1.51	8.78	1.69	9.49	1.78	0.82	2.52
H. B.....	10.56	2.64	10.19	1.70	10.42	2.40	0.71	0.23
C. S.....	5.67	1.18	7.56	1.30	7.75	1.51	5.13	5.35
W. G.....	8.89	1.22	9.33	1.41	10.36	1.69	1.40	4.20
H. V.....	7.56	1.55	6.50	0.84	7.18	1.17	3.65	1.22
L. R.....	9.81	1.79	11.61	2.43	11.31	2.67	3.58	2.80
W. T.....	4.69	0.94	4.53	1.02	5.25	1.70	0.72	1.72
W. E.....	5.28	1.54	6.14	2.36	5.50	2.24	1.83	0.49

Six of the subjects showed an increase in the average number of regressions in 12.5 percent O_2 while in 10.5 percent O_2 only four increased (Table 4). Most of the changes were slight and in many cases the variability for the number of regressions was as large or larger than the mean.

TABLE IV

MEAN NUMBER OF REGRESSIONS PER LINE

Subject	Control		12.5% O_2		10.5% O_2		$d/\sigma d$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	0.56	0.85	0.61	0.83	0.67	0.78	0.28	0.57
C. K.....	0.25	0.43	0.36	0.54	0.17	0.37	0.76	0.90
R. M.....	1.38	0.93	1.44	1.09	1.42	1.01	0.26	0.16
H. B.....	1.78	1.46	1.47	1.12	1.72	1.17	1.00	0.18
C. S.....	0.66	0.64	0.95	0.74	0.95	1.00	1.72	1.44
W. G.....	1.40	0.79	1.58	1.01	1.75	0.94	0.91	1.75
H. V.....	1.36	0.82	1.03	0.67	0.95	0.91	1.68	2.07
L. R.....	2.33	1.05	2.69	1.37	1.36	1.49	1.25	0.33
W. T.....	0.52	0.67	0.28	0.07	0.39	0.65	1.67	0.79
W. E.....	0.74	0.95	0.64	0.92	0.70	1.11	0.45	0.18

Line Divergence.—As mentioned above, this measure is the amount of convergence or divergence appearing in the eye record at the end of a line compared to what it was at the beginning of that line.

All means (Table 1, Fig. 2*B*) represent convergence on the record or a reduction in convergence of the optic axes as compared to the beginning of the line. The actual measured amount of line divergence is about .5 mm. This corresponds to an average divergence of the optic axes of slightly more than one degree during the reading of a line of print. The average amount of divergence for the group tends to remain the same in 12.5 percent O_2 and shows a slight decrease in 10.5 percent O_2, with a slight increase in the variability of the group. The individuals do not maintain their relative positions in the group for this variable as indicated by the correlations of .71 in 12.5 percent O_2 and .58 in 10.5 percent O_2 as compared with air (Table 5).

TABLE V

MEAN DIVERGENCE FROM THE BEGINNING TO THE END OF LINE (MEASURED IN MM)

Subject	Control		12.5% O_2		10.5% O_2		*d*/σ*d*	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	.51	.12	.49	.12	.51	.15	0.60	0.26
C. K.....	.38	.19	.23	.09	.36	.27	3.92	0.20
R. M.....	.98	.14	.99	.20	.98	.15	0.20	1.42
H. B.....	.29	.15	.12	.15	.33	.11	4.65	1.35
C. S.....	.48	.19	.84	.13	.59	.30	9.36	1.67
W. G.....	.17	.17	.30	.20	.10	.19	2.93	0.50
H. V.....	.53	.18	.79	.23	.71	.21	5.41	3.87
L. R.....	.67	.23	.58	.27	.35	.22	1.36	8.00
W. T.....	.44	.14	.53	.14	.49	.13	2.56	1.53
W. E.....	.48	.14	.27	.53	.22	.74	2.15	5.09

Average Binocular Adjustment During the First Fixation.—These data pertain to any shift of the optic axes relative to each other during the time that the eyes made the first fixation on the line. It was noted that the binocular adjustment on the first fixation was greater than during the succeeding ones in each line. During the fixation the adjustment was in the direction of divergence of the optic axes. The group means (Table 1, Fig. 2*B*) show that there was no change in the amount of divergence in 12.5 percent O_2 but a slight decrease in divergent adjustment in 10.5 percent O_2.

The slight increase in variability of the group was the same in both 12.5 percent and 10.5 percent oxygen. The rank order correlations are low, indicating that the various individuals in the group lost their relative positions with respect to each other in the reduced oxygen (Table 6).

Average Binocular Adjustment During All Fixations.—There is a consistent decrease (Table 1, Fig. 2*B*) in the average amount of divergence of the optic axes during all fixations. The differences between these means are significant for the group in 12.5 percent O_2 and 10.5 percent O_2. The rank

TABLE VI

AVERAGE DIVERGENCE DURING FIRST FIXATION OF EACH LINE (MEASURED IN MM)

Subject	Control		12.5% O_2		10.5% O_2		$d/\sigma d$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	.41	.07	.37	.09	.40	.11	2.00	0.64
C. K.....	.26	.11	.17	.11	.35	.10	3.07	3.39
R. M.....	.69	.17	.78	.13	.69	.10	2.43	0.01
H. B.....	.21	.09	.26	.07	.12	.08	2.64	4.09
C. S.....	.21	.12	.49	.11	.35	.16	10.23	4.28
W. G.....	.20	.14	.14	.15	.11	.10	1.68	3.19
H. V.....	.42	.13	.52	.11	.48	.19	3.57	1.51
L. R.....	.54	.20	.30	.16	.18	.15	5.01	8.22
W. T.....	.30	.10	.31	.09	.29	.08	0.61	0.42
W. E.....	.30	.07	.19	.14	.04	.17	4.01	10.17

orders are maintained with correlations of .96 and .89 for the 12.5 percent O_2 and 10.5 percent O_2 conditions in relation to the control rankings.

Eight of the ten subjects showed a decrease in divergence during fixations in 12.5 percent O_2; 5 of the decreases were significant compared with the data recorded in air (Table 7). In 10.5 percent O_2, 9 out of 10 subjects showed a decrease in divergence, all nine being significant changes.

Binocular Relations During Saccadic Movements.—The saccadic movement is generally recognized as a period of non-perception. During these saccadic movements the adjustments of the optic axes relative to each other were in the direction of convergence, and of an amount corresponding to divergence during fixations.

TABLE VII

AVERAGE DIVERGENCE DURING ALL FIXATIONS PER LINE (MEASURED IN MM)

Subject	Control		12.5% O_2		10.5% O_2		$d/\sigma d$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Control & 12.5%	Control & 10.5%
M. K.....	.145	.034	.118	.030	.121	.032	3.58	3.13
C. K.....	.118	.038	.109	.038	.123	.042	0.92	0.44
R. M.....	.257	.050	.220	.068	.196	.053	2.56	4.82
H. B.....	.058	.023	.066	.022	.045	.023	1.50	2.48
C. S.....	.173	.050	.164	.041	.139	.041	0.85	3.14
W. G.....	.079	.027	.057	.029	.064	.034	3.47	2.16
H. V.....	.196	.049	.201	.046	.147	.043	0.47	4.49
L. R.....	.130	.043	.091	.030	.062	.033	4.47	7.40
W. T.....	.238	.053	.213	.067	.173	.044	1.23	4.18
W. E.....	.183	.050	.142	.060	.087	.060	3.00	6.77

Qualitative Changes.—Representative samples of the eye movement records of six subjects taken under the three conditions, *i.e.*, control (air), 12.5 percent O_2, and 10.5 percent O_2, have been reproduced in Figs. 4 and 5 to show the qualitative changes which occurred in the low oxygen tensions as compared with the control series in air.

These photographs were selected to illustrate some of the characteristic responses which occurred in the reduced oxygen. The portion of M. K.'s record at 10.5 percent O_2, for example, shows a wavering of fixation which did not occur at any time in the control record (Fig. 4).

There were wide individual differences in the general qualitative appearance or pattern of each record. It was possible to recognize the record of each subject from the general appearances alone for most of the ten subjects. In the lower percentages of oxygen there was a tendency for the same pattern to persist and in some cases to become accentuated.

Perhaps the most regular control record is that of M. K. (Fig. 4). It is exceedingly regular with five or six fixations tending to occur in a rhythm. The tracks are quite parallel, with the exception of the first fixation. On this first fixation the adjustment can be readily seen (convergence on the record meaning actual divergence of optic axes). At 12.5 percent

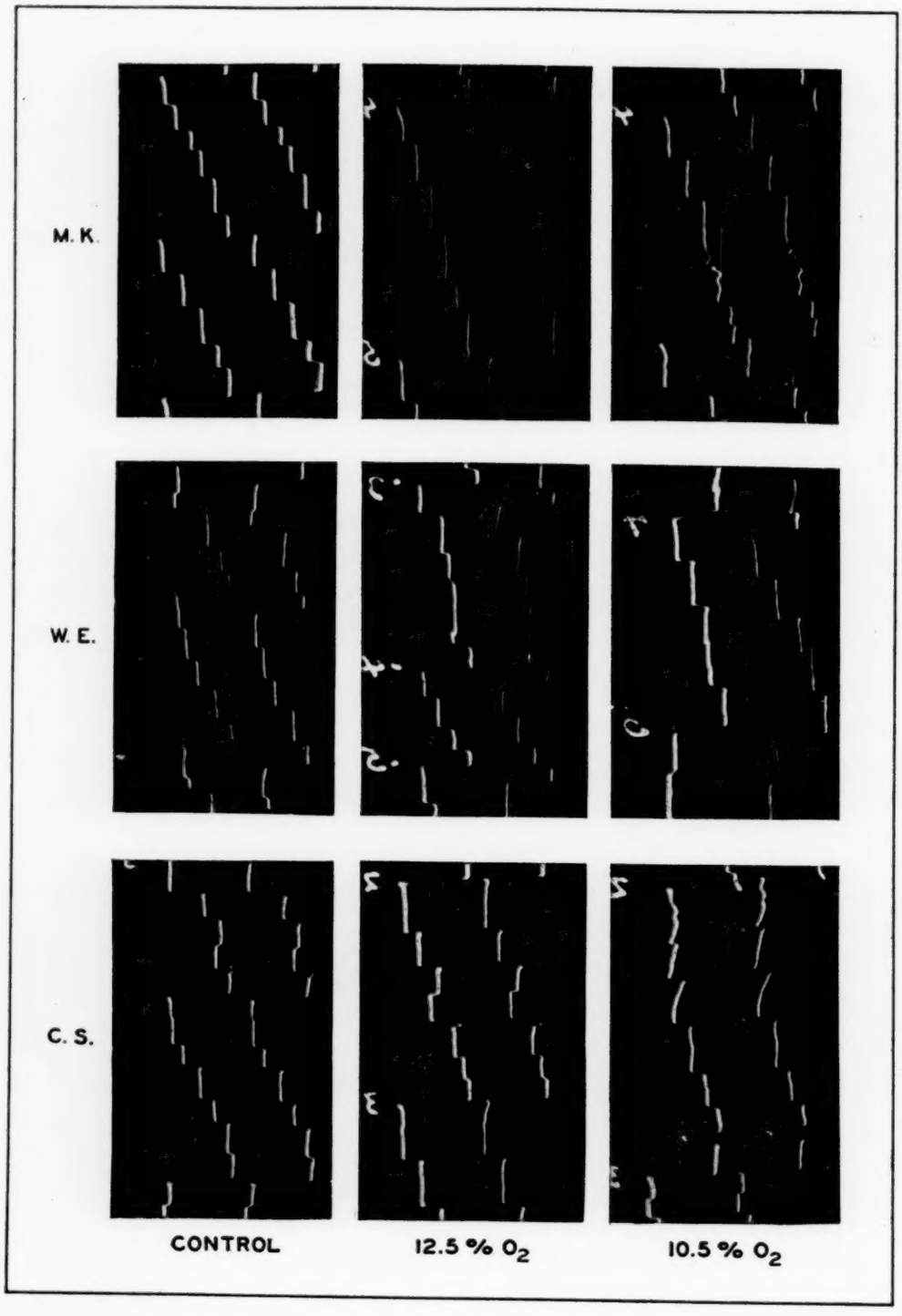

FIG. 4.

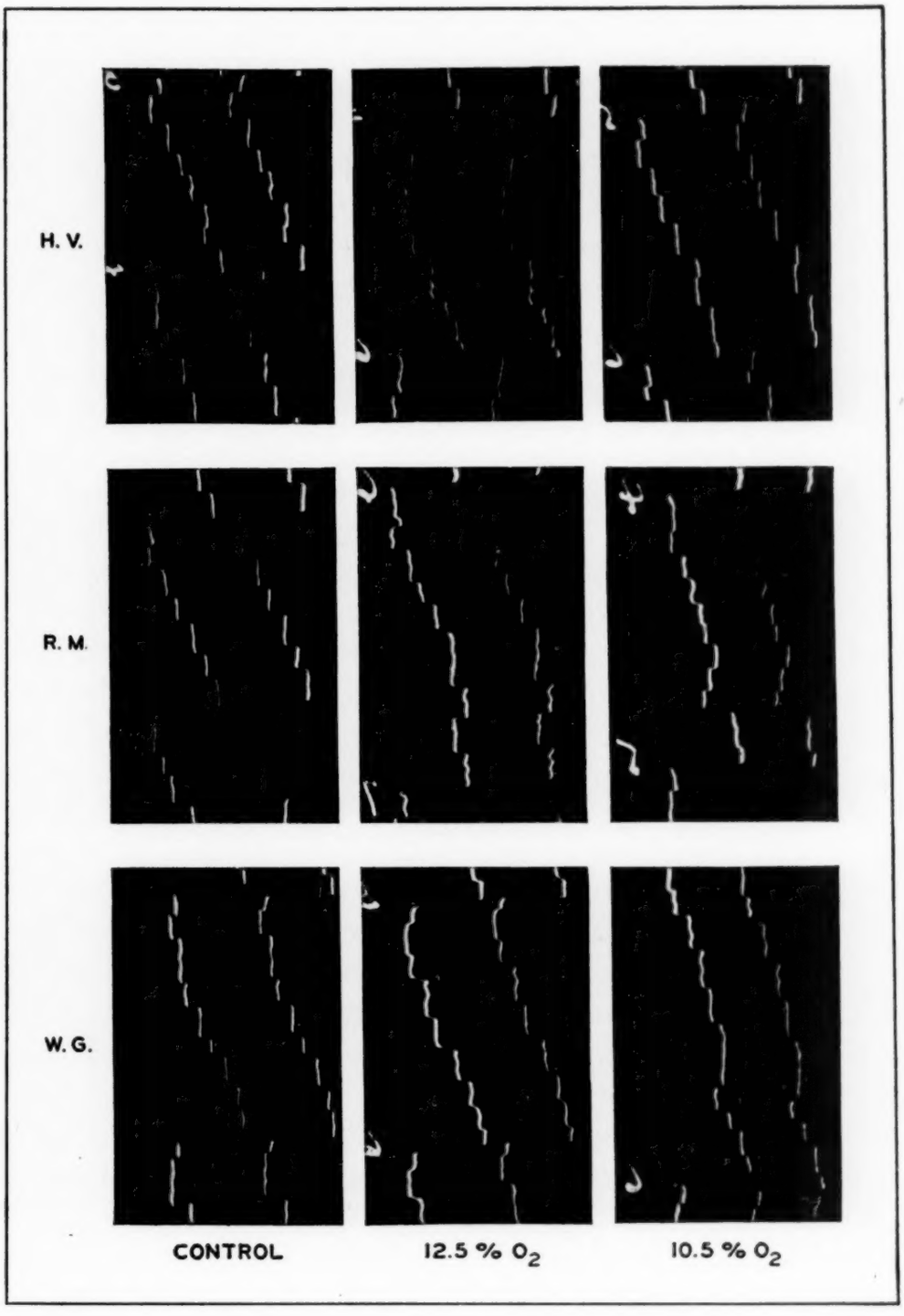

FIG. 5.

oxygen the adjustment on the first fixation is slightly greater. The fixations become longer and there seems to be a general unsteadiness. This wavering is rhythmic in all records. In 10.5 percent O_2 the fixations are also longer, and there is a general absence of clearly defined movements. A marked unsteadiness of fixation is shown in the latter part of the line.

The records of W. E. are of interest in light of the ophthalmological findings (Table 4). This subject was found to have a left hyperphoria. His reading was rapid, with few fixations. The adjustments during fixations were made with either or both eyes. The record is not very regular even in the control series. (It is in this record that an artifact mentioned above under a description of the machine must be taken into account.) At the end of the fourth fixation in 12.5 percent O_2 the right eye is seen to make a movement which is not made by the left. In the 10.5 percent O_2 record the left eye deviates. (Fourth fixation in the complete line shown.) In most of this subject's record in 10.5 percent O_2 there is a gradual increase in the distance between the tracks. Frequently, this occurred very suddenly during eye closure. These movements indicate an increasing convergence. Actually the oxygen lack unmasks the left hyperphoria, producing an approach of the corneal reflections. This deviation of the eyes could be seen by the experimenter. The subject reported that his eyes seemed to lose focus. It was difficult for the experimenter to focus the eyes on the ground glass because of unsteadiness of fixation. Comprehension was medium in 12.5 percent O_2 and poor in 10.5 percent.

The ocular photographs of subject C. S. are quite characteristic; all of his records, including the control, showing the slight 'hook' at the beginning of most fixations. The record in 12.5 percent O_2 appears to be qualitatively similar to the control one. In 10.5 percent O_2 the sliding fixations, the general unsteadiness, and the reading time are markedly increased. The ophthalmological findings indicated very poor vision and a relative central scotoma (5°) for form in the left eye (Table 8).

TABLE VIII

TABULAR SUMMARY OF OPHTHALMOLOGICAL EXAMINATION

TABLE VIII

Tabular Summary of Ophthalmological Examination

Subject	Acuity		Phorias *								Prism Vergence				P.C. R. or L.	P.P. of Accommodation		Remarks
			6 m.				25 cm.				6 m.		25 cm.					
	R.	L.	E.	×	H.R.	H.L.	E.	×	H.R.	H.L.	Div.	Conv.	Div.	Conv.		R.	L.	
M. K.	20/15–1	20/15–1		0				2			10	15	15	15	60	140/200	140/200	Interiors and exteriors normal. Fusion normal. No Rx.
C. K.	20/20	20/15		0				3–5			5	15	10	65	40	110/300	120/300	Interiors and exteriors normal. Tension normal. Slight convergence excess for near. Slight astigmatism OU. No Rx.
R. M.	20/20–1	20/20		0				0			6	20	16	35	50	120/300	120/300	Interiors and exteriors normal. Slight convergence excess for near. No Rx.
H. B.	20/30–2	20/15		1				2			6	15	10	25	80	125/300	95/300	Interiors normal. Tension normal. Migraine condition. Two chalazions upper lid. Scar upper tarsus of OD. No Rx.
C. S.	20/40–2	20/70–1		6				0			6	15	8	15	112	100/300	130/300	Conjunctiva congested. Tension normal. Left blind spot enlarged. No aniseikonia. Relative scotoma for form (5°) OS. Rx: R − .37 + .75 axis 85° = 20/20; L + .62 − .75 axis 70° = 20/40–2
W. G.	20/15	20/15		1				1			6	20	8	15	120	110/300	80/300	Interiors and tension normal. Some follicular conjunctivitis. Treatment: accommodation and convergence exercises.
H. V.	20/25–1	20/20–3		0				4–6			4	12	10	15	45	90/300	100/300	Interiors and tension normal. Mild conjunctivitis. No Rx. Treatment: accommodation and convergence exercises.
L. R.	20/20–3	20/20–3		1				0			8	15	2	12	60	120/300	120/300	Interiors and tension normal. Follicular conjunctivitis. No Rx.
W. T.	20/20–3	20/15–4		0				5			6				65			Interiors and tension normal. Conjunctivitis and sinus condition treated. Rx: R − 0.25 with − 0.75 axis 180° = 20/15–1; L − 0.62 axis 180° = 20/15–1
W. E.	20/15	20/15		1		4		8		4	5	10	2	10	135	125/300	120/300	Interiors, exteriors and tension normal. Paresis right superior rectus and spasm left inferior oblique. Blurring and diplopia after reading. No aniseikonia. Rx: R Plano 1½Δ base up; L Plano 1 Δ base down.

*Maddox Rod. E = Esophoria. × = Exophoria. H.R. = Right Hyperphoria. H.L. = Left Hyperphoria.

Large individual differences in the general appearance of the records can be seen even in the control records. Generally, there was a decrease in the precision of movements and the maintenance of fixation in reduced oxygen. In most cases nystagmoid movements crept into the records in varying amounts and tended to occur in a rhythmic manner.

Ophthalmological Examination.—The results obtained in the eye examination of each subject are shown in Table 8. Although no high correlation was found for the group between the individual tests of the ophthalmological examination and any of the measurements of the ocular photographs, we were able to interpret certain of the abnormal phenomena in the light of the examinations, as indicated above.

Correlations.—The correlations between the various measures by the rank difference method are shown below:

Reading time and Fixation/line	.94 ± .02
Reading time and Regression/line	.92 ± .03
Reading time and Adjustment per line	− .38 ± .18
Fixation/line and Regression/line	.88 ± .05
Fixations and Adjustment all fixations	− .50 ± .16
Regressions/line and Adjustment all fixations	− .30 ± .19

The correlation is very high (.94) between rapid reading time and few fixations. It is also high (.92) between fast reading time and few regressions. There is, however, little or no correlation between time and the other measures. A rapid rate of reading, few fixations and few regressions tend (as evidenced by the negative coefficients) to be associated with the larger adjustments during fixation. These correlations range from − .30 to − .50.

Comprehension.—The comprehension ratings for the group show a decline when the subjects breathed reduced concentrations of oxygen. None of the subjects fell into the poor comprehension class until the effects of the anoxemia became marked in 10.5 percent O_2.

COMPREHENSION RATINGS

	Control	12.5% O_2	10.5% O_2
Good	80%	50%	30%
Medium	20%	50%	50%
Poor	0%	0%	20%

Pulse and Blood Pressure.—The changes in pulse and blood pressure were only slight, indicating that the circulatory changes were probably not extreme enough to account for the observed deviations in ocular movements. The results of these tests are shown below:

		Control	12.5% O_2	10.5% O_2	Control	5% CO_2
Pulse....	15 min.	74	79	89	79	81
	30 min.	73	78	83	79	78
	1 hr.	74	74	83	77	79
B.P.....	15 min.	116/69	112/62	114/57	115/71	118/70
	30 min.	117/69	108/64	112/55	113/70	117/75
	1 hr.	115/70	109/64	110/59	113/70	117/74

In order to show the degree of acclimatization during each experimental period in 12.5 percent O_2 and 10.5 percent O_2 the group means are presented graphically in Fig. 3 at 15, 30, and 60 minutes after exposure to the reduced oxygen. The reading time in 12.5 percent O_2 (Fig. 3*A*) increased at the end of 30 minutes and decreased at the end of one hour. In 10.5 percent O_2 the mean reading time was not only increased but also the average subject was slower at the end of each hour compared with the beginning, *i.e.*, acclimatization did not take place. Likewise, Fig. 3*B* shows that an average decrease in divergence took place in 12.5 percent O_2, gradually returning toward the control conditions at the end of the hour. No such adjustment took place in 10.5 percent O_2. Similar findings were obtained with the other measurements. The average subject, therefore, appeared to acclimatize to 12.5 percent O_2 (13,500 ft) as indicated by an improvement in reading capacity as the experiment progressed. In 10.5 percent O_2 (18,500 ft), however, the average subject showed a steady decrease in efficiency of ocular performance toward the end of the hour.

In an attempt to compare the effects of reduced oxygen with excess carbon dioxide four subjects were tested in the chamber with a normal supply of O_2 and 5 percent CO_2. The results are shown in Table 1. The differences between the control and excess CO_2 series were not significant, but

indicate several minor and consistent changes. There is a slight decrease in the average time and number of fixations with no change in the number of regressions. The data also indicate a consistent decrease in the amount of divergence and the individual and group variability. The qualitative changes showed marked irregularities. In some cases, however, it was difficult to control the head movements due to the increased rate and depth of breathing associated with the excess CO_2. An increase in CO_2 above 5 percent may have shown more consistent and significant changes.

Discussion of Results

The results of this experiment indicate that the effects of oxygen deprivation on eye movements in reading increase the time and frequency of fixations and reduce the adequacy of the ocular adjustments, *i.e.*, the normal amount of adjustment during each fixation is diminished. The decrease in the precision of the ocular movements was apparent in 12.5 percent O_2, corresponding to 13,500 feet and quite marked in 10.5 percent O_2 (18,000 ft). These effects may be ascribed to the diminished amount of oxygen being delivered to the nervous tissue in general but it is difficult to attribute the reaction to cortical more than to sub-cortical areas of the brain. Gildea and Cobb (*cf.* **8** and **12**) believe that the cortex is considerably more sensitive to anoxemia than the brain stem and spinal cord. Gellhorn and Spiesman (**11**) have reported that the oxygen lack must be more severe to effect nystagmus movements as compared with auditory acuity and after-images and they attribute the difference to the location of the important neurons or synapses involved, *i.e.*, sub-cortical in the case of nystagmus and cortical in vision and audition. The reactions studied in this experiment no doubt involved both cortical and sub-cortical areas of the brain. The fact that there was a slowing down, as well as a significant loss of motor coördination in ocular movements in relatively slight degrees of oxygen lack (with only a slight impairment of comprehension) tends to indicate that the motor areas of the brain were affected as soon and as acutely as the sensory and associative areas.

The fact that there was no increase in the average number of regressions suggests that either the organism failed to adjust, or that the degree of comprehension was not seriously affected, *i.e.*, the average subject did not regress more often to get the meaning of the words, while reading under conditions of anoxemia. In 10.5 percent O_2, however, the degree of comprehension was significantly diminished. Since it has been shown by McFarland (**11**) that the capacity for immediate memory is markedly reduced in the more acute stages of oxygen deprivation this was no doubt a contributing factor. The absence of a significant increase in number of regressions, as well as a decrease in the accuracy and amount of fixation in the ocular movements, suggests that O_2 lack, the same as Scofield (**17**) observed in the effects of alcohol on nystagmus, causes the organism to more readily abandon the attempt to adjust than it does under normal conditions. The loss of efficiency, therefore, cannot be attributed to psychological changes or variations in conscious effort but to the direct effects of the oxygen lack on the nervous tissue and consequently on nervous functions.

The changes observed in the ocular movements while reading in reduced oxygen can hardly be attributed to a diminution in attention since such variations would be apparent in the eye records. Vernon (**21**) observed that lack of attention lead to the most rapid reading and the most regular movements. Nor can these variations be attributed to gross alterations in circulation. The changes in pulse rate and blood pressure were not extreme in any of the healthy, young individuals who served as subjects. It is well-known that anoxemia produces vasodilatation and increased blood flow, however this vasomotor adjustment is not sufficient to completely counteract the effects of the anoxemia. It seems more reasonable to attribute the changes to the diminished excitability of the nervous tissue itself due to the diminished arterial oxygen saturation and consequently the lack of oxygen being delivered to the nervous tissue. This is substantiated by the fact that the eye movements returned to normal conditions upon the administration of oxygen.

The measured decrease in divergence in reduced oxygen represents an actual decrease in the ocular adjustments while reading. If the amount of adjustment under control conditions be considered adequate for reading under those conditions, then they should be of the same magnitude to be adequate in low oxygen. There was, however, a significant decrease in the amount of adjustment under low oxygen, indicating that the adjustment was not adequate.

The divergence during fixations and convergence during saccadic movements are apparently not related to the heterophorias in the eye examinations. There is, however, a very close relationship between vergence and accommodation, *i.e.*, with changes in accommodation there are concomitant changes in vergence. An increase in accommodative effort results in an increase in convergence, with a resulting relaxation in convergence with a diminution of accommodative effort. This relaxation can occur while the eyes are marking time, *i.e.*, waiting for perception. The result may be a divergence with respect to the point of regard for near vision, which is said to be common (Verrijp, *cf.* 4).

The relaxation seems to stay within such bounds that the images on the two retinas are on corresponding fusional areas resulting in a maintenance of single vision. Ames, *et al.* (1) report that it is not necessary for images to fall on corresponding points for fusion, but that fusion is maintained as long as images fall on corresponding fusional areas. This may account, therefore, for the subjective maintenance of clear vision during the variation of vergence while reading.

From the data presented in this experiment, it appears that there is no clear relationship between the direction and the amount of adjustment and the findings on the eye examinations (particularly the acuity, the heterophorias, fusional reserve, and the near points of accommodation and convergence, *cf.* Table 8). Nine of the ten subjects had a muscle balance within normal limits and hence correlations could hardly be expected. The hyperphoria of one subject, however, was quite apparent on the photographs in reduced oxygen. There was also an accentuation of defects noted in the control records of one other subject. It might be said that the eye

disorder does not show on the eye record until the abnormality becomes somewhat accentuated or the oxygen deprivation fairly acute. At the present time we are extending the investigation with special reference to patients with accentuated motor anomalies (**14**).

In comparing the findings of this investigation of the effects of oxygen lack with those of other investigators, it is quite apparent that deterioration was manifested with milder degrees of anoxemia, *i.e.*, at partial pressures of oxygen corresponding to lower altitudes. Dunlap and Bagby (*cf.* **11**), for example, in summarizing the psychological studies on pilots during the World War, state that they were unable to observe significant changes until immediately before collapse or general psychomotor deterioration (9–7 percent or 22,000–27,000 ft). In those experiments with the rebreather the tests were, for the most part, of short duration (20–30 minutes) and the initial effects were concealed by the subject's exertion of great effort. In this study of eye movements the psychological factors such as variations in effort, or knowledge of results, would seem to be insignificant. The test, therefore, has proved to be a sensitive one in detecting the early effects of oxygen deprivation and hence it may prove to be of practical value as an objective measure of the effects of high altitude in aviation.

Summary

1. A significant increase was found for the group in reading time at oxygen percentages of 12.5 and 10.5, corresponding to altitudes of 13,500 and 18,000 feet. The increase in the number of fixations did not become significant until 10.5 percent oxygen. There was no change in the average number of regressions for the group. The increase in time, without a corresponding increase in regressions, may indicate a failure of the organism to adjust and an impairment of cortical and sub-cortical nervous centers.

2. A definite tendency was found for divergence of the visual axes during fixations and for convergence during the saccadic movements in all conditions. This was perhaps related to changes in convergence due to variations in accommodative effort.

3. A significant decrease was found in the amount of divergence occurring during fixations. This may represent an inadequacy of ocular adjustments while reading in low oxygen.

4. In general qualitative characteristics, the eye movements showed a marked diminution of precision in diminished oxygen, with the appearance of rhythmical nystagmoid movements. There was a tendency for abnormalities to be accentuated in reduced oxygen.

5. The findings on the eye examination did not correlate with the other data in this experiment for the group as a whole. It was important, however, in the interpretation of the records of two individuals whose vision was not normal.

6. A decrease in the comprehension of the reading material was noted with diminished oxygen tensions. This may represent an impairment of memory for immediate recall.

7. The decrease in efficiency in ocular movements under anoxemia is attributed to the diminished amount of oxygen being delivered to the nervous tissue. Changes in pulse and blood pressure were slight under diminished oxygen as compared to the control conditions.

8. In 5 percent carbon dioxide there was a tendency for a decrease in reading time, fixations, and amount of divergence of visual axes, but the changes were not statistically significant.

9. The data indicated that the subjects tended to acclimatize in 12.5 percent oxygen by the end of the hour. No such adjustment was made at 10.5 percent O_2.

10. It is suggested that the average reading time per line and the adjustment during fixations may be sensitive measures of the early effects of anoxemia. The test seems to be free from the effects of increased effort, which in many tests masks the impairment.

(Manuscript received May 15, 1937)

References

1. Ames, A., Ogle, K. N., and Gliddon, G. H., Corresponding retinal points, the horopter, and size and shape of ocular images, *J. Opt. Soc. Amer.*, 1932, **22**, Nos. 10 and 11, 537–631.

2. Anonymous, Rhythm Reading. *The Ophthalmograph*, Amer. Opt. Co., Chicago, 1935, pp. 35.
3. Berens, C., Present ophthalmologic standard for commercial aviation in the United States, *J. Aviation Med.*, 1932, **3**, No. 2, 55–101.
4. Berens, C. (editor), *The Eye and its Diseases*, Saunders. Philadelphia, 1936, pp. 1254.
5. Dodge, R., Five types of eye movement in the horizontal median plane of the field of regard, *Amer. J. Physiol.*, 1903, **8**, 307–329.
6. Dodge, R., An experimental study of visual fixation, *Psychol. Rev.*, 1907, **8**, No. 4.
7. Dodge, R. and Benedict, F. G., *Psychological effects of alcohol*, Carnegie Inst. 1915, Pub. No. 232, pp. 281.
8. Gellhorn, E. and Spiesman, I. G., The influence of hyperpnea and of variations of O_2 and CO_2 tension in the inspired air upon after images, *Amer. J. Physiol.*, 1935, **112**, 620; upon nystagmus, *Ibid.*, 1935, **112**, 662; upon visual intensity discrimination, *Ibid.*, 1936, **115**, 679.
9. Goldmann, H. and Schubert, G., Untersuchungen über das Gesichtsfeld bei herabgesetztem Sauerstoffdruck der Augensluft, *Arch. f. Augenk.*, 1933, **107**, 216.
10. Javal, E., Essai sur la physiologie de la lecture, *Annales d'oculistique*. 1878, **79**, 97–117; 240–274; 1878, **80**, 135–147; 1879, **81**, 61–73; 157–162; 1879, **82**, 72–81; 159–170.
11. McFarland, R. A., The psychological effects of oxygen deprivation (anoxemia) on human behavior, *Archiv. of Psychol.*, 1932, No. **145**, pp. 135.
12. McFarland, R. A., Psycho-physiological studies at high altitudes in the Andes. Part II. Sensory and motor responses during acclimatization, *J. Comp. Psychol.*, 1937, **23**, 226 ff. Part III. Mental and psychosomatic responses during acclimatization. *Ibid.*, 1937, **23**, No. 3.
13. McFarland, R. A. and Barach, A. L., The relationship between alcoholic intoxication and anoxemia, *Amer. J. Med. Sci.*, 1936, **192**, 186. *Cf.* also McFarland, R. A. and Forbes, W. H., The metabolism of alcohol in man at high altitudes. *Human Biol.*, 1936, **8**, 387–398.
14. McFarland, R. A., Knehr, C. A., and Berens, C., The effects of anoxemia on ocular movements while reading. With special reference to abnormalities in clinical patients (to be published), *Proc. Assoc. Res. in Ophthalm.*, 1937.
15. Schmidt, W. A., An experimental study in the psychology of reading, *Suppl. Educ. Monog.*, 1917, No. **2**.
16. Taylor, E. A., *Controlled Reading*, Univ. Chicago Press. Chicago, 1937, pp. 345.
17. Sc ofield, C. F., The effects of mild doses of alcohol and caffeine on optic nystagmus, *Psychol. Monog.*, 1936, **47**, No. 2, 217–241.
18. Tinker, M. A., Reliability and validity of eye movement measures of reading, *J. Exper. Psychol.*, 1936, **19**, 732–746.
19. Tinker, M. A., Eye movements in reading, *J. Educ. Res.*, 1936, **30**, No. 4, 1–37.
20. Vernon, M. D., *The Experimental Study of Reading*, Cambridge, Cambridge Univ. Press, 1931, pp. xv–190.
21. Vernon, M. D., *The movements of the eyes in reading*, Special Report Series, No. 148, Medical Research Council, His Majesty's Stationery Office, London, 1930, pp. 45.
22. Wilmer, W. H. and Berens, C., Medical studies in aviation. V. The effect of altitude on ocular functions, *J. Amer. Med. Assoc.*, 1918, **71**, 17, 1394–1398.

THE EFFECT OF INDUCED MUSCULAR TENSION UPON VARIOUS PHASES OF THE LEARNING PROCESS

BY JAMES C. STAUFFACHER *

Introduction

Although a great deal of research has been reported on the interdependence of various mental activities and the motor processes, a relatively small amount of work has been reported on the inter-relationship between learning and muscular tension. Investigations which have been done upon this problem may be divided into two main sorts of inquiries: (1) What is the effect of learning upon muscular tonicity? (2) What is the effect of changes in muscular tension upon the efficiency of learning?

Prominent among studies following the first line of inquiry are the following. McTeer (**11**) found that punishment in the form of an electric shock tended to produce greater tension in the inactive hand during the course of mirror drawing activity. Clites (**5**) used the three water-dipping problems found at the superior adult level of the Binet test and found an increase in the winking rate accompanying this type of activity. Book (**4**) found that when a subject learning to typewrite made a special effort to overcome errors, he used extra force in hitting the keys. Stroud (**14**) found a correlation between the speed of learning and the amount of tension as measured by the pressure on the stylus in human maze learning. Ghisseli (**9**) worked with two multiple choice visual motor speed tasks and found that tension was maintained at a constant level during learning and that once the task was learned the tension markedly decreased.

The other line of approach used in studying the interdependence of learning and muscular tonicity consists of changing the degree of muscle tension and noting the change

* From the Psychological laboratory of the University of Chicago.

in efficiency of learning. For example, Reed (**13**) reported that learning scores were affected by the posture assumed at the time of learning (sitting being more advantageous than standing) but that recall was not affected by a shift in posture at the time of recall. Since the sitting position is the more customarily assumed during work of this type, the influence of habit would probably give this effect regardless of the muscular tensions involved. On the other hand, Bills (**1**) induced muscular tension by the use of the hand dynamometer and found that the efficiency of learning was increased. Stroud (**14**) reported that the rate of learning the stylus maze was materially increased when muscular tension was artificially induced in the hands of the subject.

The results of these experiments indicate in general that there is a correlation between a certain degree of muscular tension and efficiency of learning. This conclusion clears the way for further studies of the effects of muscular tension. The study here reported deals with the investigation of several of these problems.

I. What is the effect of varying degrees of induced muscular tension upon the rate of learning?
II. What is the effect of muscular tension induced during the original learning upon retroactive interference?
III. What is the effect of induced muscular tension introduced during relearning upon the efficiency of recall and relearning?

Experiment I

The Degree of Induced Muscular Tension and the Rate of Learning

Problem.—Several studies have shown that a certain amount of induced muscular tension is facilitative to learning. These conclusions would lead one to think that perhaps different degrees of induced tension introduced during learning might produce different effects upon the efficiency of that process. This experiment is an investigation of this problem. Stated in other words the question reads, "Is there a certain degree of tension which will give a maximum amount of

facilitation in learning and if so, approximately how much is this?"

Apparatus.—Tension was induced by weights suspended over pulleys which were attached to the far side of a table. The amount of tension could be accurately varied by combining the various weights. The subjects gripped the stirrup-handles, lifted the weights off the floor and supported them. This procedure was successful in minimizing the distraction, since no effort or attention was required from the subject in order to keep the induced tension constant. In using the heavier of the weights, it was found necessary to screw the table securely to the floor, to insure its stability.

An electrically driven memory drum which advanced every two seconds was used to expose the material to be learned. This material consisted of four lists of fifteen, three letter nonsense syllables.

Auxilliary handles were attached to the movable arms of two Smedley spring dynamometers which were clamped securely to the top of the table. The subject could then pull on these in the same manner that they were required to do in order to support the weights. This device was used to measure the relative strength of the subjects.

Subjects.—Forty graduate and undergraduate students in the Department of Psychology served as subjects in this experiment.

Procedure.—A preliminary test was given each subject in order that the amount of weight used for the induction of tonus could be made proportional to the strength of the individual. The subjects were instructed to exert their maximum pull on the two auxiliary handles of the dynamometers for one minute. The strength of the pull at the end of this period was taken as the maximum pull. The work was carried on under four conditions: (1) No induced tension, (2) One-quarter of the total tested pull, (3) One-half of the total tested, and (4) Three-quarters of the total tested. Every subject worked under all four conditions, one condition on each day (usually consecutive). The order of the various conditions (*A*, *B*, *C*, and *D*) and the syllable lists used (1, 2, 3, and 4) were systematically varied from subject to subject, in order to counterbalance practice and the relation of list to condition as shown in Table I.

The instructions to the subjects were as follows: "I have placed a list of fifteen, three letter nonsense syllables on this revolving drum. [The nature of these syllables

TABLE I

ORDER OF CONDITIONS AND LISTS USED IN EXPERIMENT I

Subject	First Day	Second Day	Third Day	Fourth Day
1	*A*(1)	*B*(2)	*C*(3)	*D*(4)
2	*B*(1)	*C*(2)	*D*(3)	*A*(4)
3	*C*(1)	*D*(2)	*A*(3)	*B*(4)
4	*D*(1)	*A*(2)	*B*(3)	*C*(4)
5	Same as subject 1.			
etc.				

was explained to any subjects not familiar with them.] The drum revolves intermittently, so that each syllable appears in the window for two seconds. I want you to learn as many of the syllables as you can. Learn as many as possible on the first revolution of the drum and when the XX's appear, you are to spell out the first syllable. When the first syllable appears, try to spell the second one and so on through the list, each syllable acting as the stimulus for you to spell the one that follows. You will be given six trials and you are to spell as many as possible on each trial after the first."

The instructions for the tonus conditions were: "I want you to support these weights while you are learning. Grasp the handles and pull until your elbows are even with the side of your body." A demonstration was given if the subject did not understand what was expected of him.

Results.—Two criteria were used in measuring the efficiency of the learning under the various conditions of muscle tension. First, a comparison was made of the number of syllables correctly spelled on the last trial under the four different conditions. This is shown in Table II. Condition

TABLE II

THE EFFECT OF DEGREE OF TENSION UPON THE NUMBER OF SYLLABLES ANTICIPATED CORRECTLY ON THE LAST TRIAL (AVERAGES FOR FORTY SUBJECTS)

Condition	Amount of Induced Tension	Average No. of Syllables Anticipated Correctly on Last Trial
A	0	4.8 ± 0.25
B	¼ total tested	5.3 ± 0.22
C	½ total tested	5.6 ± 0.25
D	¾ total tested	5.3 ± 0.26

Conditions	R.O. Correl.	Difference	Difference / P.E. Diff.
A and *B*	.478	.5 ± 0.24	2.1
A and *C*	.648	.8 ± 0.19	4.4
A and *D*	.511	.5 ± 0.25	2.0
B and *C*	.665	.3 ± 0.19	1.7
B and *D*	.693	0	—
C and *D*	.631	.3 ± 0.22	1.4

C differed most from Condition *A* (the normal condition), the difference being .8 ± .19. This difference was 4.4 times its P.E. It will be remembered that Condition *C* is that condition where the tension is one-half the maximum tested load. All conditions of induced tension used, however, were slightly facilitative when compared with the condition of normal tension. These differences are only about two times their P.E.'s and are of questionable reliability.

A comparison was also made of the most syllables anticipated correctly on any one trial under the four conditions as shown in Table III. The use of this criterion was suggested by the observed fact that a subject might name very few or even no syllables on a given trial due to his having gotten a poor start or for other reasons. If this happened on the last trial, then the number anticipated during this trial would be a poor criterion of the amount of learning accomplished. The use of the most number of syllables correctly anticipated on any one trial does not depend upon performance during any given trial and so is not subject to this criticism. Again

TABLE III

THE EFFECT OF DEGREE OF TENSION UPON THE LARGEST NUMBER OF SYLLABLES ANTICIPATED CORRECTLY ON ANY ONE TRIAL (AVERAGES FOR FORTY SUBJECTS)

Condition	Amount of Induced Tension	Average No. of Syllables Anticipated Correctly on All Trials
A	0	5.2 ± 0.23
B	¼ total tested	5.6 ± 0.21
C	½ total tested	6.3 ± 0.23
D	¾ total tested	5.4 ± 0.25

Conditions	R.O. Correl.	Difference	Difference / P.E. Diff.
A and B	.477	.4 ± 0.23	1.7
A and *C*	.629	1.1 ± 0.19	5.8
A and *D*	.557	.2 ± 0.23	.9
B and *C*	.633	.7 ± 0.18	3.9
B and *D*	.682	.2 ± 0.19	.1
C and *D*	.697	.9 ± 0.18	5.0

Condition *C* showed the greatest difference from the normal condition, this difference being 1.1 ± 0.19 (5.8 times it's P.E.). The other tension conditions were also slightly facilitative when this criterion was used.

An attempt was made to compare the performance of the best and the poorest learners under the four conditions of muscle tension. The individuals selected did not divide themselves equally into the four sequences of conditions, however, so that there was no way of eliminating the effect of practice in the comparison of individual differences. Even

if this were possible, it is very improbable that an individual score would mean a great deal since such a small sample of each subject's performance was used. Only the group averages are of significance.

This experiment indicates, then, that there is a certain amount of muscular tonus which, when induced in the subject, will cause facilitation of the learning. This degree of tension is somewhere in the neighborhood of one-half the individual's maximum, when tested in the manner used in this experiment. Lesser or greater amounts of induced muscular tension gave little or no facilitation, but this result was not statistically reliable.

Experiment II

Muscular Tension and Retroactive Interference

Problem.—The evidence from the previous experiment has shown that a certain kind and amount of induced muscular tension results in an increase in the efficiency of learning. The next question to be considered is whether material learned under these conditions of induced tonus is as well integrated, *i.e.*, as resistant to disrupting influences, as material learned under 'normal' conditions. (This does not follow from the demonstrated fact that induced tension facilitates the rate of learning.) Specifically, the purpose of this experiment is to test the degree to which material learned under induced tension is susceptible to retroactive interference, as compared with material learned without induced muscular tension. Retroactive interference was produced by introducing an interpolated list between the learning and recall of the original list.

Apparatus.—The tonus was induced in the same manner as that used in Experiment I, the amount of weight used being ten pounds per hand for the men and six pounds per hand for the women. This is approximately the optimum amount as determined by the first experiment.

The material learned consisted of six lists of fifteen, common, four-letter words.

Subjects.—As in the first experiment forty graduate and undergraduate students in psychology served as subjects. Approximately half of these subjects had already served in the previous experiment.

Procedure.—The conditions of the experiment were as shown in Table IV.

In the conditions where interpolated learning was done, this followed immediately upon the completion of the original learning. There was a twenty-four hour interval

between learning and relearning in each case. The conditions and lists used under each condition were varied systematically to counterbalance the effect of practice and possible inequality in the associative value of the lists of words.

TABLE IV

CONDITIONS OF WORK FOR EXPERIMENT II

Day	Original Learning	Interposed Learning	Recall and Relearning
First	Normal	Normal	Normal
Second	Tension	Normal	Normal
Third	Tension	None interpolated	Normal
Fourth	Normal	None interpolated	Normal

The directions to the subjects were: This is an experiment in learning. I have typed a list of fifteen, common, four-letter words and placed them on this drum. The drum revolves intermittently, so that each word appears in the window for two seconds. I want you to memorize this list of words. Learn all that you can on the first revolution of the drum and when the XX's appear, you are to say the first word. When the first word appears, say the second word and so on through the list. Each word is to act as the stimulus for you to say the following word. The word must be spoken before it appears in the window to be counted. The trials will continue until you are successful in correctly naming all the words on one revolution of the drum. Now, are you ready?

If the condition called for induced tension, the subject was told, "While you are learning, I want you to support the weights with both hands, like this." The experimenter would demonstrate the method of holding the weights, then, with the right hand since he was seated at the left of the subject.

If interpolated learning was required, the windows on the drum were adjusted immediately upon the completion of the original learning and the subject was told, "Now I would like to have you learn another list without the use of the weights." The subjects were not informed before this of the fact that two lists were to be learned on that particular day.

Just before the subject left he was admonished as follows, You are not to think of the words on this list (or these lists, as the case might be) between now and tomorrow. If you find that they are running through your mind, actively repress them by thinking of something else.

At the time of recall the instructions to the subjects were: "I want you to recall the list (or the first list) of words which you learned yesterday. Name all of them that you can on the first trial and then relearn them. You are not to use the weights. Ready?"

Results.—The learning scores under the two conditions confirm the results of the first experiment. The average number of trials to learn are shown in Table VII. It will be seen that in both cases learning was more efficient under the tension condition. The average number of trials

required to learn under the normal condition (with and without interpolation) was 14.9 ± 0.63 and for tension, 13.3 ± 0.51. This is only a difference of 1.6 ± 0.81. One reason for this small amount of difference undoubtedly is the fact that the same amount of weight was used for all the subjects and this was too great for some individuals and not large enough for others.

It is easy to understand the size of the P.E.'s when one remembers that there was no preliminary practice given to the subjects during this experiment. The counter-balanced order was relied upon to eliminate the effect of practice upon the averages. Since there was a great deal of practice improvement during the experiment (some of the subjects cutting the number of trials required to learn down to one-third of the first day's record) this introduced a large degree of variation in the figures and was naturally reflected in the corresponding size of the P.E.'s.

In another study by Bills and Stauffacher (**2**), in which the mental work was the solution of 'Detectograms' (miniature mystery stories) and the tonus was induced in the same way and in approximately the same amounts, it was found that the good performers, as measured by both the time and error criteria, were made less efficient by the introduction of tonus and the poor performers were facilitated in their work.

It was rather difficult to analyze the data of the present experiment in the same manner since there was a great deal of practice effect during the course of the learning. The counter-balanced order would take care of this for the group but it is difficult to appraise an individual's score under the two conditions. The best way to make this analysis seemed to be to take an average of the learning scores of all the subjects who learned under tension on each day; then to compute the mean of these averages; and finally, to compare this mean with the mean learning score under the normal condition after it had been computed in the same manner. Fourteen subjects were selected as good learners on the basis of the total trials to learn all the original learning lists, and fourteen poor learners were selected on the same basis. The

average number of trials to learn as determined by the method just described is shown in Table V.

The only significant effect of tension was in the case of the poor learners who were able to learn in eight trials less during induced tension. This means that only fourteen of the forty subjects were actually facilitated in their learning by the induction of the tension.

There are several possible explanations for this fact. It is true that most of the good learners were those who had served as subjects in other learning experiments and thus might be called 'practiced' subjects. Since this previous practice in learning had taken place under the so-called normal condition, the induction of tension was a distracting change in the condition of the learning for these subjects.

TABLE V

COMPARISON OF GOOD AND POOR LEARNERS UNDER NORMAL AND TENSION CONDITIONS (TRIALS TO LEARN)

	Normal	Tension	Difference
Good performers (14).....	19.5	20.9	1.4 (N)
Poor performers (14).....	40.5	32.6	7.9 (T)
All others (12).....	29.2	27.0	2.2 (T)

Another explanation might be found in the level of tonus maintained under the normal condition. These good performers were working at a high degree of efficiency. Since efficiency in mental work is correlated with a certain degree of muscular tonus, the induced tension may have placed these subjects at a level of muscular tonicity which exceeded the facilitating amount and caused a corresponding fall in the efficiency of the learning.

Another possible explanation of this phenomenon might be suggested. Perhaps the good subjects are those who work up to the limit of their ability and the poor learners, those who do not work up to the limit of their possibilities. The tension may then be a stimulus which operates in such a manner as to make those subjects work at a higher level of efficiency. Induced tension would then be effective only

with those subjects who are working below the level of their highest potential efficiency.

These data were further analysed by dividing the subjects into two groups: one, those who had had no previous practice in learning by the anticipation method, and the other, those who had had such practice. The results of an analysis made by the same method as that used in comparing the good and the poor performers, in average number of trials to learn are as shown in Table VI.

TABLE VI

COMPARISON OF PRACTICED AND NON-PRACTICED SUBJECTS UNDER NORMAL AND TENSION CONDITIONS (TRIALS TO LEARN)

	Number of Subjects	Normal	Tension	Difference
Non-practiced Subjects				
Good learners	8	20.2	19.4	.8 (T)
Poor learners	8	50.1	32.6	17.5 (T)
All others	8	32.1	27.7	4.4 (T)
Practiced Subjects				
Good learners	6	18.7	22.3	3.6 (N)
Poor learners	6	36.9	32.6	4.3 (T)
All others	4	26.3	26.2	.1 (N)

This table shows that the principal effect of tension occurred in the case of the eight, unpracticed, poor learners. This effect was almost entirely concealed when the total group of forty subjects was considered.

The degree to which material learned under induced tension is susceptible to retroactive interference as compared with material learned without induced tension is shown in a comparison of the recall, savings and relearning scores.

The recall scores are given in Table VII in terms of the average numbers of words correctly anticipated during the first relearning trial. The amount of decrement in recall of material learned under the normal condition which results from the interpolated work is $5.98 - 4.33$ or 1.65. The amount of decrement in recall of material learned under the

induced tension condition which results from similar interpolated learning is 5.63 − 4.48 or 1.15. There is, then, .5 or 30 percent less decrement in recall after the interpolation when tension is induced at the time of the original learning.

TABLE VII

THE EFFECT OF INDUCED TENSION UPON THE REINSTATEMENT OF MATERIAL LEARNED UNDER 'NORMAL' TENSION

	Without Interpolation			With Interpolation		
	Learning	Recall	Relearning	Learning	Recall	Relearning
Normal...	14.2±0.75	5.98±0.38	5.53±0.30	15.5±0.78	4.33±0.36	6.4±0.25
Tension...	13.4±0.52	5.63±0.37	4.75±0.27	13.2±0.66	4.48±0.31	5.1±0.29
Difference.	.8±0.91	.35±0.53	.78±0.41	2.3±1.02	.15±0.48	1.3±0.39

The savings scores, using the formula

$$S = 1 - \frac{\text{trials to relearn}}{\text{trials to learn}} \times 100,$$

are as follows:

	Normal	Tension
Without interpolation.....	61.2%	64.5%
With interpolation.....	58.7%	61.8%
Difference.....	2.5%	2.7%

Thus according to the savings scores there is practically no difference between the two conditions. One would not expect to find a difference here, since the same subjects and lists were used under the two conditions, unless the induced tension produced more resistance to retroactive interference than it did facilitation of learning. Since the tension condition did facilitate learning and there was no difference in savings scores, one is led to the conclusion that the resistance to retroactive interference was not increased more than the efficiency of learning was facilitated by the induced tension.

The difference between the two conditions is more evident, however, when the relearning scores are compared. Relearning scores are determined by two factors: the amount of retention and the learning ability of the subject. Since the same lists were used under both conditions, and the

subjects were identical, relearning scores, in this case, are a measure only of the amount of retention.

The difference between the normal and the tension conditions when there was no interpolated learning is .78 ± 0.41 and when there was interpolated learning the difference becomes 1.3 ± 0.39. This difference is made more striking when the correlations between the learning and relearning of the forty subjects under the two conditions are taken into consideration. The rank order correlations, using the formula $\rho = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$, are as follows:

Between Learning under Normal and Tension (No Interpolation)	.362
Between Relearning under Normal and Tension (No Interpolation)	.123
Between Learning under Normal and Tension (With Interpolation)	.409
Between Relearning under Normal and Tension (With Interpolation)	.348

The differences between learning and relearning under the two conditions and their P.E.'s with the correlations taken into consideration by use of the formula

$$PE_d = \sqrt{(PE_1)^2 + (PE_2)^2 - 2_\rho PE_1 \cdot PE_2}$$

are as follows:

Without Interpolation		With Interpolation	
Learning	Relearning	Learning	Relearning
.8±0.74	.78±0.38	2.3±0.79	1.3±0.31

It will be noticed that whereas the difference in relearning was .78 ± 0.38 (critical ratio of 2) when there was no interpolated learning, the difference in relearning with interpolation became 1.3 ± 0.31 (critical ratio of 4). This is certainly a significant difference in favor of the tension condition.

An attempt was made to compare the retroactive effects in the case of the good and poor performers by the use of the same analysis which was used for learning. No differentiation could be made on this basis.

This experiment has indicated that there is a slight tendency for the resistance to retroactive interference to be increased by the induction of a certain degree of tension during the original learning. This degree of tension was approximately that amount which had proved to facilitate

learning. There is also an indication that poor learners benefit from the induced muscular tension in terms of increased efficiency in learning whereas the good learners are inhibited. Among the poor learners, those who had had no previous practice in the anticipation method of learning seemed to benefit much more than the practiced poor learners.

Experiment III

Muscular Tension and Relearning

Problem.—There are many factors which complicate the picture and obscure the effect of muscular tonus on the mental processes. The first two experiments have demonstrated a few of those. The effect of induced muscular tension on learning depends upon its degree, the level of efficiency of the subject and his experience with the task involved. The second experiment also indicated that learning is 'stamped in' in a more stable fashion if that learning has been accomplished under a certain amount of induced muscular tension. The problem to be dealt with in this experiment is the determination of the effect of induced muscular tension introduced during recall and relearning upon the efficiency of reinstatement.

Apparatus and Subjects.—The tonus was induced in the same manner as in Experiment II. The material used was also of the same type, two different lists of fifteen, common, four-letter words being used in this case. As in the other experiments, forty graduate and under-graduate students in psychology served as subjects.

Procedure.—The conditions of the experiment were as follows:

Learning	Relearning
Normal	Normal
Normal	Tonus

The position of the induced tension during relearning was varied so that any differences due to practice effect or to the difficulty of the lists were counter-balanced.

The instructions given to the subjects before and after the learning were the same as in the previous experiment, as were the instructions at the time of relearning. The interval between learning and relearning was twenty-four hours.

Results.—The results of this experiment are given in Table VIII. Condition I is that condition in which relearning was done under normal tension and Condition II is the condition in which relearning was done under induced tension.

It will be seen that none of the differences are significant. The difference between trials taken to relearn and the number of words recalled on the first trial is, in both cases, about equal to its probable error. The savings scores were also computed and they show practically no difference between the two conditions. The savings score for Condition I is 74.5 percent and for Condition II, 74.2 percent.

TABLE VIII

THE EFFECT OF TENSION INDUCED DURING REINSTATEMENT UPON THE EFFICIENCY OF REINSTATEMENT

	Learning (Trials)	Recall (Words Named)	Relearning (Trials)
Condition I.....	17.8±0.80	8.23±0.31	4.6±0.34
Condition II.....	19.9±1.71	8.78±0.35	5.1±0.69
Difference.....	2.1±1.81	.55±0.47	.5±0.77

It will be noticed, however, that in Condition II the induction of tension during reinstatement caused the recall and relearning to be done under a condition which was not similar to that under which it was learned. There are, then, two factors which influence the efficiency of the reinstatement under Condition II: the induced tension, and the change in the condition. If reinstatement under Condition II had proven to be more efficient than under the other (normal) condition, in spite of the detrimental effect of the change of conditions, then the results would have been unambiguous. As it was, however, another experiment seemed to be indicated, the conditions of which would be as follows:

	Learning	Relearning
Condition I.....	Tension	Normal
Condition II.....	Tension	Tension

This experiment was conducted using twenty subjects and the same apparatus and procedure as that used in the previous experiment, except for the change in the conditions which has been described. The results are given in Table IX.

It will be seen that there is no significant difference between the recall or the relearning under the two conditions.

TABLE IX

THE EFFECT OF TENSION INDUCED DURING REINSTATEMENT UPON THE EFFICIENCY OF REINSTATEMENT OF MATERIAL LEARNED UNDER TENSION

	Learning (Trials)	Recall (Words Named)	Relearning (Trials)
Condition I	17.9±2.57	5.9 ±0.72	5.61±0.42
Condition II	18.1±2.84	6.15±0.49	5.9 ±0.56
Difference	.2±3.84	.25±0.87	.29±0.70

An analysis was made of the effect of the tension condition upon the recall and relearning of the good and poor learners, in the same way in which the analysis of learning was made in Experiment II. The results are shown in Table X. It

TABLE X

COMPARISON OF RECALL AND RELEARNING OF GOOD AND POOR LEARNERS UNDER NORMAL AND TENSION CONDITIONS OF REINSTATEMENT

	No. of Subjects	Normal	Tension	Difference
		Recall		
Poor learners	10	5.5	6.2	.7 (T)
Good learners	10	6.3	6.1	.2 (N)
		Relearning		
Poor learners	10	6.9	6.3	.6 (T)
Good learners	10	4.5	5.5	1.0 (N)

will be seen that there is an indication that poor learners were facilitated in both recall and relearning by the induction of the tension, and that the good learners were inhibited. The most striking result occurs in the case of the relearning of the good learners when the inhibitory effect of the tension is over 20 percent.

This experiment, then, shows no over-all effect in recall or relearning from the induction of tension at the time of reinstatement. There is an indication, however, that poor learners are facilitated and good learners inhibited both in recall and relearning by the tension condition. In this connection, it will be remembered that Experiment II showed

that tension induced during learning facilitated the learning of the poor learners, but inhibited the learning efficiency of the good learners.

Conclusions

1. This study corroborates previous investigations in finding that a certain amount of induced muscular tension facilitates learning. (Experiments I and II.)

2. Degrees of induced muscular tension which are smaller or greater than the facilitating amount give little or no facilitation of learning. (Experiment I.)

3. Although there is a slight indication that material learned under a degree of induced muscular tension which is facilitative to learning is stamped-in more firmly and consequently there is less decrement in recall and relearning after interpolated learning than is present when the original learning takes place under the 'normal' conditions of tension, these results are not reliable and no final conclusion can be drawn from them. (Experiment II.)

4. Poor learners benefit from induced muscular tension in terms of increased efficiency in learning, whereas good learners do not. (Experiment II.)

5. Among the poor learners, those who have had no previous practice in the anticipation method of learning benefit much more than the practiced poor performers. (Experiment II.)

6. Recall and relearning efficiency are not significantly affected by the introduction of induced muscular tension during the reinstatement when the whole group of subjects is considered. (Experiment III.)

7. There is an indication, however, that poor learners may be facilitated and good learners inhibited in both recall and relearning by the induction of tension during reinstatement. (Experiment III.)

8. These experiments have indicated that there is a facilitating effect of induced muscular tension but that it is modified by many factors such as the amount of tension used, the length of the task, the level of performance of the subject,

and his past experience (practice in the task). It is difficult to demonstrate the effect of tonus with reliability unless these factors are controlled.

Discussion and Interpretation

Muscular tension does not produce its effect upon the efficiency of mental performance in the same way that many so-called 'distractors' do, *i.e.*, by drawing the subject's attention away from the task so that an increase in effort (which actually results in facilitation of the performance) is necessary to overcome this effect. Morgan (**12**) found that when a sound distraction was long continued it gradually ceased to exert a distracting influence. He reports that there was an increased tension throughout the period of the distraction, however, even after the distraction had become ineffective in causing a work decrement. Perhaps the presence of the tension is responsible for the increased efficiency which Morgan reported to have resulted from the distraction. Bills (**1**) found that the increase in efficiency gained by the use of tension was enhanced with practice. It is evident that the effect of the tension is not like that of a distracting sound since in that case it would gradually become less effective with continued practice.

Ford (**6**), working with a problem solving situation during severe sound distraction, found that as 'automatization' became more and more complete tonus as measured by a pressure pencil became progressively less. Instead of the tonus acting as a distractor, again it seemed to be one of the factors which was effective in overcoming the effect of the distraction exerted by other influences.

The theory of G. L. Freeman (**8**) offers a more plausible explanation of the way in which muscular tonus operates in producing an effect upon the mental processes. He assumes that cortical centers involved in mental work have high thresholds. When these thresholds are lowered the centers will react to stimuli which would otherwise be inadequate. This is accomplished by the spread of proprioceptive impulses which travel to the cortex from lower centers, especially those

involved in muscular contraction. If this proprioceptive stimulation falls below a minimum level, cortical neurones become incapable of adequate response. The work of Jacobson on relaxation substantiates the theory at this point. Above a certain maximum, tension becomes an inhibitor of precise neural integration. The facilitative effect is only within limits, which vary presumptively from individual to individual. The establishment of a facilitative level of tension in Experiment I confirms this portion of the theory.

Experiment I showed that there is a certain amount of tonus which, when induced in the subject, will cause a facilitation of the learning. The lesser or greater amounts of induced muscular tension which were used gave little or no facilitation. The continuous line in Fig. 1 shows these

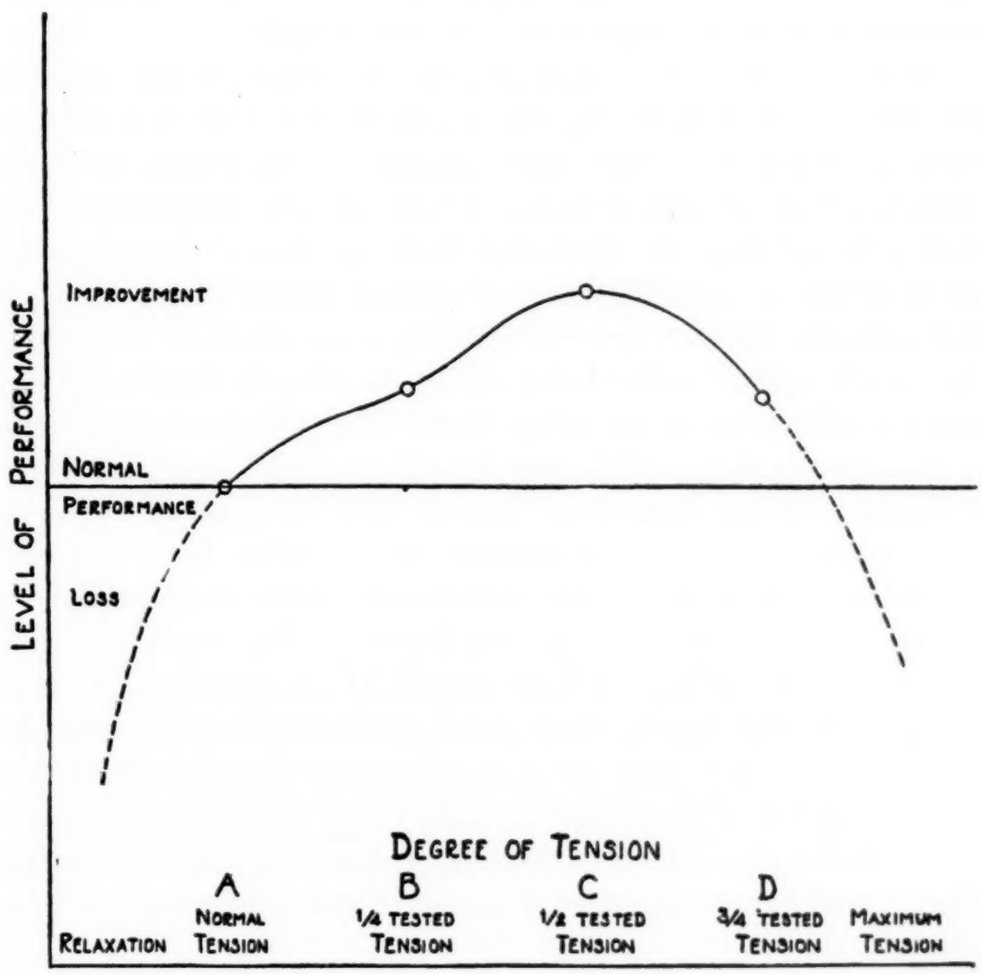

FIG. 1. Relation between degree of tension and levels of performance.

results. It will be noticed that the levels of performance of Conditions *B* and *D* were placed approximately midway between the level of performance given by Condition *C* and that of the normal condition. These were the actual findings but the differences between the normal condition and Conditions *B* and *D* were only about two times their P.E. Because of the questionable reliability of these differences, the statement has been made that they gave "little or no" facilitation.

The studies of Jacobson indicate that if a subject is relaxed to a much greater degree than normal, mental work is possible only at a much lower level of efficiency than under normal conditions. Freeman (**8**) showed that a high degree of induced tension will cause a decrement in the efficiency of mental work. The results found by these two investigators are indicated in the figure by the dotted lines.

Since there is a certain degree of tonus for the highest efficiency of mental work, one would expect that if a subject were working at a very high degree of efficiency, a corresponding high degree of tonus would already be present and that any increase in muscular tonicity would prove to be detrimental to the efficiency of mental work. This is what one actually finds. Several studies, such as that of Stroud (**14**) with stylus maze learning, have shown that with increased efficiency of learning there is an incidental increase in muscle tension. Bills and Stauffacher (**2**), working on a problem solving situation, found that the most efficient performers worked at a higher level under the 'normal' condition and the poorer performers were more efficient during the induced tension condition. This result is confirmed by the findings of Experiment II on learning. One is led to the conclusion that good performers were already working at a very high level of efficiency (correlated with a certain degree of muscular tonicity) and that the increased tonus placed their level of muscle tonus over the facilitating degree and therefore caused a decreased efficiency of performance.

This is one of the factors which are probably responsible

for the large amount of individual differences which appear in many studies of the effect of varying amounts of muscle tension. The neglect of such factors is illustrated by the study of the effect of varying amounts of muscular tension which was reported by Block (**3**). She found that 22 pounds tension induced in both arms was optimal for half of her eight subjects when speed was the criterion of efficiency and for three-quarters of her subjects when accuracy was the criterion. It is surprising that these percentages were as large as they were since one would expect the facilitating degree of tension to vary with the individual. In other words, a degree of induced tonus which would be facilitative for one individual might be actually detrimental for another. Then in any study of the effect of induced muscular tension in which a fixed amount of tension was used, one would expect to find subjects who were not facilitated by that amount of tension unless careful selection of the subjects were practiced.

The results of this study tend to confirm Washburn's 'Motor Theory of Consciousness' which asserts that the motor processes are not only accompaniments of consciousness but that they play a direct causal rôle in consciousness. Muscular processes do affect the efficiency of mental work. The subject reacts as a whole and his mental, sensory and motor processes can only be separated ideationally, since each influences the others and is in turn influenced by them.

(Manuscript received December 14, 1936)

References

1. Bills, A. G., The influence of muscular tension on the efficiency of mental work, *Amer. J. Psychol.*, 1927, **38**, 227–251.
2. Bills, A. G., and Stauffacher, J. C., "The influence of voluntarily induced tension on rational problem solving." J. of Psychol., 1937.
3. Block, H., The influence of muscular exertion upon mental performance, *Arch. of Psychol.*, 1936, No. **202**.
4. Book W., F., *The Psychology of Skill.* New York: Gregg, 1925.
5. Clites, M. S., Certain somatic activities in relatoin to successful and unsuccessful problem solving, *J. Exper. Psychol.*, 1935, **18**, 708–724.
6. Ford, A., Attention automatization. An investigation of the transitional nature of mind. *Amer. J. Psychol.*, 1929, **41**, 1–32.
7. Freeman, G. L., Mental activity and the muscular processes, *Psychol. Rev.*, 1931, **38**, 428–449.

8. FREEMAN, G. L., AND KATZOFF, E. F., Muscular tensions and irritability, *Amer. J. Psychol.*, 1932, **44**, 789–792.
9. GHISELLI, E., Changes in neuro-muscular tension accompanying the performance of a learning problem involving constant choice time, *J. Exper. Psychol.*, 1936, **19**, 91–99.
10. JACOBSON, E., Electrical measurements of neuro-muscular states during mental activities.
    I. Imagination of movement involving skeletal muscle, *Amer. J. Physiol.*, 1930, **91**, 567–608.
    II. Imagination and recollection of various muscular acts, *Amer. J. Physiol.*, 1930, **94**, 22–34.
    III. Visual imagination and recollection, *Amer. J. Physiol.*, 1930, **95**, 694–702.
    IV. Evidence of contraction of specific muscles during imagination, *Amer. J. Physiol.*, 1930, **95**, 702–712.
    V. Variation of specific muscles contracting during imagination, *Amer. J. Physiol.*, 1931, **96**, 115–121.
    VI. A note on mental activities concerning an amputated limb, *Amer. J. Physiol.*, 1931, **96**, 122–125.
11. MCTEER, W., Changes in grip tension following electric shock in mirror tracing, *J. Exper. Psychol.*, 1933, **16**, 735–742.
12. MORGAN, J. J. B., The overcoming of distraction and other resistances, *Arch. of Psychol.*, 1916, **35**, 84.
13. REED, H. J., The influence of a change of conditions upon the amount recalled, *J. Exper. Psychol.*, 1931, **14**, 632–649.
14. STROUD, J., The rôle of muscle tensions in stylus maze learning, *J. Exper. Psychol.*, 1931, **14**, 606–631.

THE GENERALIZATION OF CONDITIONED RESPONSES

III. Extinction, Spontaneous Recovery, and Disinhibition of Conditioned and of Generalized Responses

BY CARL IVER HOVLAND

Department of Psychology, Institute of Human Relations, Yale University[1]

In the writer's work on generalization of conditioned responses with varying intensities of sound stimuli (4) there appeared in the data a suggestion that conditioned responses extinguish much more slowly than generalized responses.[2] This finding is in accord with implicit assumptions of Pavlov (7) and is especially interesting in relation to the recent work of Brogden and Culler (2) which shows much more rapid extinction of responses based upon 'higher-order' motivation than of those based upon the primary unconditioned stimulus. Also in the study of the writer on generalization of tonal frequencies (3) there were indications that the increased differentiation which takes place after differential reinforcement is reduced by the passage of time. This suggests the spontaneous recovery of extinguished generalized responses. But the techniques used in the above-named investigations are not well adapted to the study of these phenomena in analytic detail, since the large number of stimuli required for the determination of the gradient of generalization prevents closely spaced repetitions of identical stimuli. The technique of the present investigation obviates this difficulty by using

[1] The writer wishes to express his deep appreciation to Professor Clark L. Hull in whose unit of the Institute of Human Relations this research was conducted.

This is the third of a series of four articles presented as a dissertation to the faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Psychology, 1936. (See also 3, 4.)

[2] The term "generalized responses" is used to refer to responses to stimuli which were not used in the process of conditioning but are similar to the conditioned stimulus. See footnote 2 in a previous article of the writer (3).

only two stimuli, chosen to represent opposite extremes of the intensity scale. The slope of the gradient of generalization can thus be determined very rapidly, and successive presentations of identical stimuli can be spaced more closely. With this technique even such a transient phenomenon as 'disinhibition' (Pavlov, **7**, pp. 62 ff.) is amenable to study.

Apparatus

A detailed description of the apparatus may be found in an earlier article (3). The Tarchanoff phenomenon of the skin, measured by a galvanometric circuit, was used as the index of response. The unconditioned stimulus was an electric shock. Conditioned stimuli were vibro-tactual and auditory. The conditioned stimuli were presented for 400 milliseconds. After a pause of 90 milliseconds the shock was given for 75 milliseconds.

Extinction and Spontaneous Recovery of Conditioned and of Generalized Tactile Conditioned Responses

Procedure.—For the tactile conditioned stimuli the vibrator developed by Bass and Hull (1) was employed. This instrument consists simply of a light-weight solenoid operated on alternating current to produce 120 vibrations per second. The base of the stimulator was cemented to the subject's left arm one inch above the wrist with the rubber-like adhesive, *Grippit.* The alternating current for the operation of the vibrator was supplied by the output of a Lionel type T transformer. For the intense stimulation the vibrator was activated by 11.3 volts and 2.7 amperes, while for the weak intensity 5.3 volts and 1.8 amperes were supplied.

Forty subjects, paid by the hour, were used. As previously shown by the writer (5) the size of the conditioned galvanic response is markedly affected by the intensity of the stimulus, so half of the subjects were conditioned to the weak intensity of the vibratory stimuli, and half to the strong. The data from the two groups were then pooled. By this procedure the extent of irradiation was measured independently of the absolute magnitude of the reactions to the two intensities.

The experimental program on the first day for the group of subjects conditioned to the weak intensity of the vibrator was as follows:

Step

I. Test response to weak and strong intensity of vibrator six times each.
II. Reinforce weak intensity of vibrator with shock stimuli sixteen times.
III. Test response to weak and strong intensity four times each.

On the second day:

IV. Test response to weak and strong intensity of vibrator.

An identical program was followed with the second group of twenty subjects, except that the reinforcements were always given to the strong intensity of the vibrator.

Comparison of the responses during steps III and IV will give an indication of the amount of spontaneous recovery taking place after 24 hours. The potentiometer circuit used for getting relatively uniformly sized responses from subject to subject was maintained in the same position as on the previous day for each subject.

Since extinction during the test trials is inevitable, a counterbalanced practice order of *a b b a* type was employed, so that tests were made to the weak and strong intensities of stimulation in arrangements such as the following:

Conditioning to strong intensity of vibrator; testing with stimuli in order: weak, strong, strong, weak, strong, weak, weak, strong.

Conditioning to weak intensity of vibrator; testing with stimuli in order: weak, strong, strong, weak, strong, weak, weak, strong.

Conditioning to strong intensity of vibrator; testing with stimuli in order: strong, weak, weak, strong, weak, strong, strong, weak.

When the entire set of data was pooled, the same number of responses to each intensity of stimulus was obtained in each test position.

Results.—The last two responses of the two groups of subjects to the weak and strong vibratory stimuli before conditioning indicate no reliable differences between the group conditioned to the weak and the group conditioned to the strong intensity of stimuli.

	Stimuli	
	Strong	Weak
Group I (Conditioned to weak vibrator)	3.13	2.13
Group II (Conditioned to strong vibrator)	2.56	1.77

The average values of the last four reinforced responses for the two groups were:

Group I	20.87
Group II	23.08

The magnitudes of galvanic response on each of the four test trials of the conditioned responses are given in Table 1. The same data for the generalized responses are presented in Table 2 for each of the subjects. These results are summarized in Table 3, which gives the mean values for the entire groups.

Graphical representation of the relationships among the various responses is found in Fig. 1. In Fig. 2 the magnitudes of response on successive test trials are shown for the conditioned and generalized responses.

The much more rapid extinction of the generalized than of the conditioned responses is apparent. If the sharp decline in the magnitude of the generalized response shows the typical characteristics of extinction, the passage of time should bring about an augmentation of the generalized responses. This

TABLE 1

MAGNITUDE OF GALVANIC RESPONSE (IN MM.) TO VIBRATORY CONDITIONED STIMULUS ON SUCCESSIVE TEST TRIALS FOLLOWING 16 REINFORCEMENTS

Group Conditioned to Strong Intensity

Subject	Test Trials			
	1	2	3	4
I	21.6	10.7	10.1	13.6
II	16.1	12.7	14.1	12.3
III	20.9	15.9	6.1	17.0
IV	12.1	14.1	12.9	23.7
V	15.4	13.9	11.7	6.8
VI	16.8	14.5	16.6	12.1
VII	13.9	10.4	11.5	9.8
VIII	4.3	9.8	7.8	10.3
IX	16.7	23.4	22.7	25.2
X	14.3	19.1	15.6	19.4
XI	17.4	18.4	20.9	22.2
XII	19.8	23.6	23.1	23.6
XIII	17.3	13.7	16.1	12.9
XIV	18.2	16.4	14.3	13.7
XV	14.6	23.3	14.1	20.9
XVI	23.9	17.3	23.8	11.1
XVII	18.7	13.0	21.7	12.6
XVIII	26.7	19.9	20.1	10.4
XIX	13.1	13.0	8.7	16.4
XX	10.3	9.1	16.4	7.8
Mean	16.60	15.61	15.41	15.09
σ dist.	4.76	4.30	5.11	5.42

Group Conditioned to Weak Intensity

I	18.2	15.8	22.6	18.0
II	14.8	4.9	11.3	11.1
III	20.8	20.0	7.8	7.8
IV	11.7	18.1	17.3	13.2
V	13.6	24.2	21.2	24.2
VI	9.8	13.7	14.1	9.9
VII	10.2	9.8	4.3	5.4
VIII	11.4	13.2	12.0	10.1
IX	7.3	6.7	9.2	5.6
X	11.4	10.4	8.6	16.9
XI	16.3	16.3	12.4	13.6
XII	13.5	10.8	13.2	10.9
XIII	17.6	16.5	14.5	13.2
XIV	11.7	12.3	10.1	17.3
XV	19.6	9.9	11.4	13.2
XVI	16.3	10.7	9.3	14.5
XVII	6.8	9.9	13.6	10.4
XVIII	10.7	11.1	5.6	4.3
XIX	9.4	10.4	10.1	9.9
XX	12.8	7.9	7.6	8.1
Mean	13.19	12.63	11.81	11.88
σ dist.	3.98	4.92	4.62	4.92

TABLE 2

MAGNITUDES OF GALVANIC RESPONSE (IN MM.) TO INTENSITY OF VIBRATORY STIMULUS NOT USED IN CONDITIONING PROCESS, ON SUCCESSIVE TRIALS FOLLOWING 16 REINFORCEMENTS

Subject	Test Trials			
	1	2	3	4
Generalized Responses to Weak Intensity				
I	10.2	8.1	9.7	9.8
II	10.6	16.5	10.1	2.6
III	20.8	15.6	7.2	3.0
IV	10.8	13.8	15.2	14.1
V	3.6	9.4	10.2	7.6
VI	22.4	10.1	9.4	10.1
VII	17.4	14.5	13.6	9.4
VIII	13.8	12.1	13.8	7.9
IX	15.5	17.1	13.4	10.3
X	13.7	9.6	10.4	8.1
XI	8.9	10.4	8.1	6.3
XII	14.0	11.1	9.2	10.1
XIII	13.1	18.9	10.3	15.6
XIV	20.3	19.2	18.7	20.3
XV	18.4	17.4	15.0	16.1
XVI	9.6	11.1	14.1	18.4
XVII	10.4	16.9	12.3	5.8
XVIII	11.3	9.8	7.6	8.3
XIX	12.7	10.1	8.9	6.7
XX	10.3	8.1	9.1	7.6
Mean	13.39	12.99	11.31	9.90
σ dist.	4.69	3.49	3.01	4.88
Generalized Responses to Strong Intensity				
I	24.2	21.8	20.4	15.6
II	26.9	15.0	11.9	1.1
III	29.8	22.0	17.8	16.0
IV	12.5	17.0	10.0	10.0
V	30.0	20.8	22.6	23.0
VI	17.3	13.8	10.4	9.6
VII	12.8	9.4	8.4	10.1
VIII	13.5	10.2	6.7	9.4
IX	9.9	11.4	10.2	10.0
X	23.2	10.7	9.4	17.5
XI	18.7	15.4	11.8	9.4
XII	16.5	13.9	10.4	6.7
XIII	19.8	17.7	9.3	13.4
XIV	26.0	9.9	6.7	8.4
XV	20.2	16.6	5.8	7.6
XVI	13.1	8.1	10.2	11.6
XVII	8.8	3.9	6.4	5.8
XVIII	12.2	9.5	11.0	9.7
XIX	10.7	10.9	8.9	10.3
XX	14.2	10.3	9.5	8.4
Mean	18.01	13.41	10.89	10.68
σ dist.	6.65	4.83	4.38	4.62

TABLE 3

AVERAGE MAGNITUDES OF GALVANIC RESPONSE (IN MM.) TO CONDITIONED AND ALLIED INTENSITIES OF VIBRATORY STIMULI ON SUCCESSIVE TEST TRIALS FOLLOWING 16 REINFORCEMENTS

Based upon 40 subjects: half conditioned to strong and half to weak intensity of vibrator.

	Test Trials			
	1	2	3	4
Group Conditioned to Strong Intensity of Vibrator				
Conditioned responses	16.60	15.61	15.41	15.09
Generalized responses	13.39	12.99	11.31	9.90
Group Conditioned to Weak Intensity of Vibrator				
Conditioned responses	13.19	12.63	11.81	11.88
Generalized responses	18.01	13.41	10.89	10.68
Combined Groups				
Conditioned responses	14.89	14.12	13.61	13.48
Generalized responses	15.70	13.20	11.10	10.29

possibility was tested experimentally. The average values of the conditioned and the generalized responses on the last two test trials of the first day were as follows:

Conditioned	13.48	$P.E._M$	0.58
Generalized	10.29	$P.E._M$	0.50
		$r =$	+0.26
Diff.	3.19	$P.E._{diff.}$	0.66
Critical ratio		$\frac{Diff.}{P.E._{diff.}}$	4.84

After 24 hours the values of the two types of responses became:

Conditioned	12.32	$P.E._M$	0.55
Generalized	12.05	$P.E._M$	0.51
		$r =$	+0.15
Diff.	.27	$P.E._{diff.}$	0.69
Critical ratio		$\frac{Diff.}{P.E._{diff.}}$	0.39

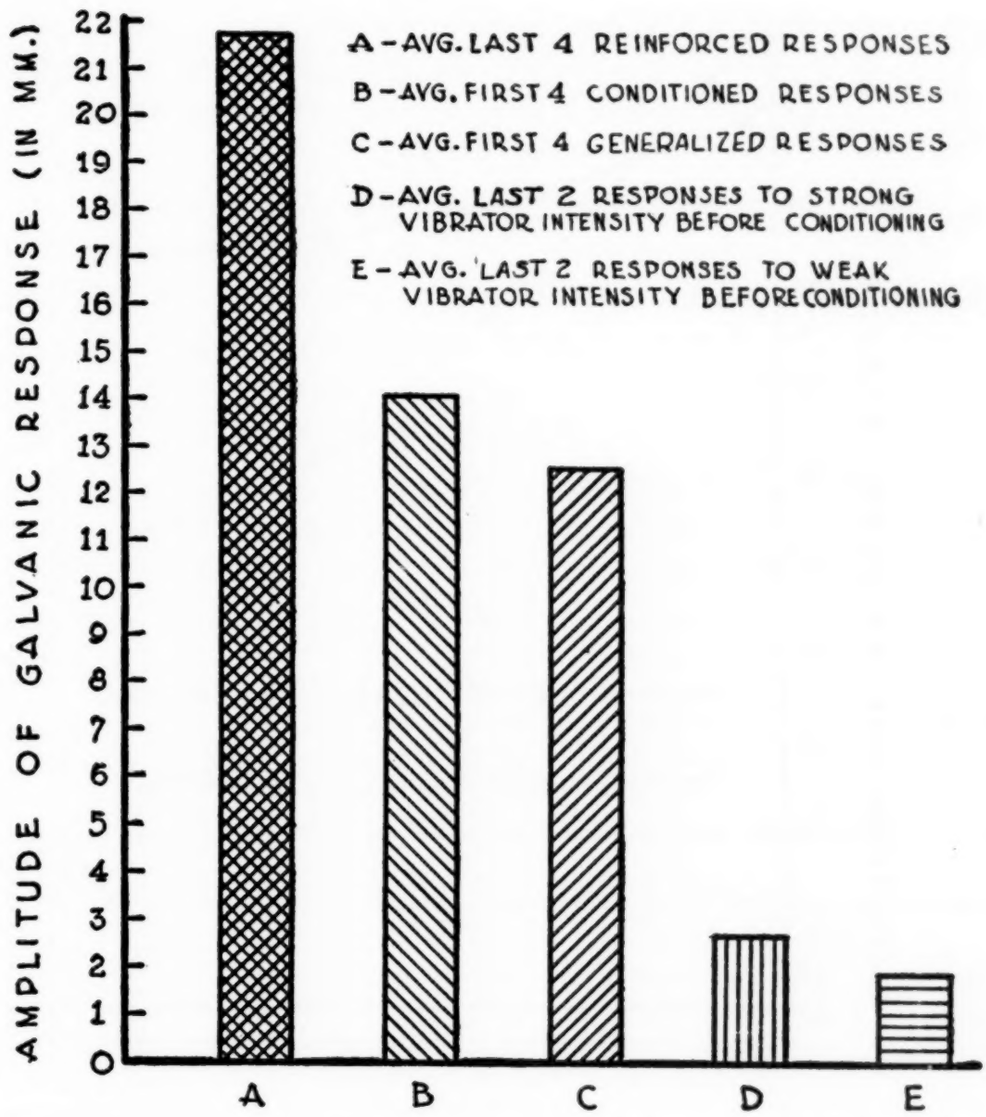

FIG. 1. Composite column diagram showing magnitudes of reinforced, unconditioned, conditioned, and generalized galvanic skin responses.

The difference in the magnitudes of conditioned and of generalized responses is no longer significant. The slope of the generalization gradient in the first case (response to conditioned stimulus divided by response to generalized stimulus) is 1.30, while after 24 hours it declines to 1.02.

From the hypothesis that the decreased generalization after the test trials is due to the rapid extinction of the generalized responses, there would follow also the deduction that 'disinhibition,' if introduced at a time when some differentiation had been established, should result in an

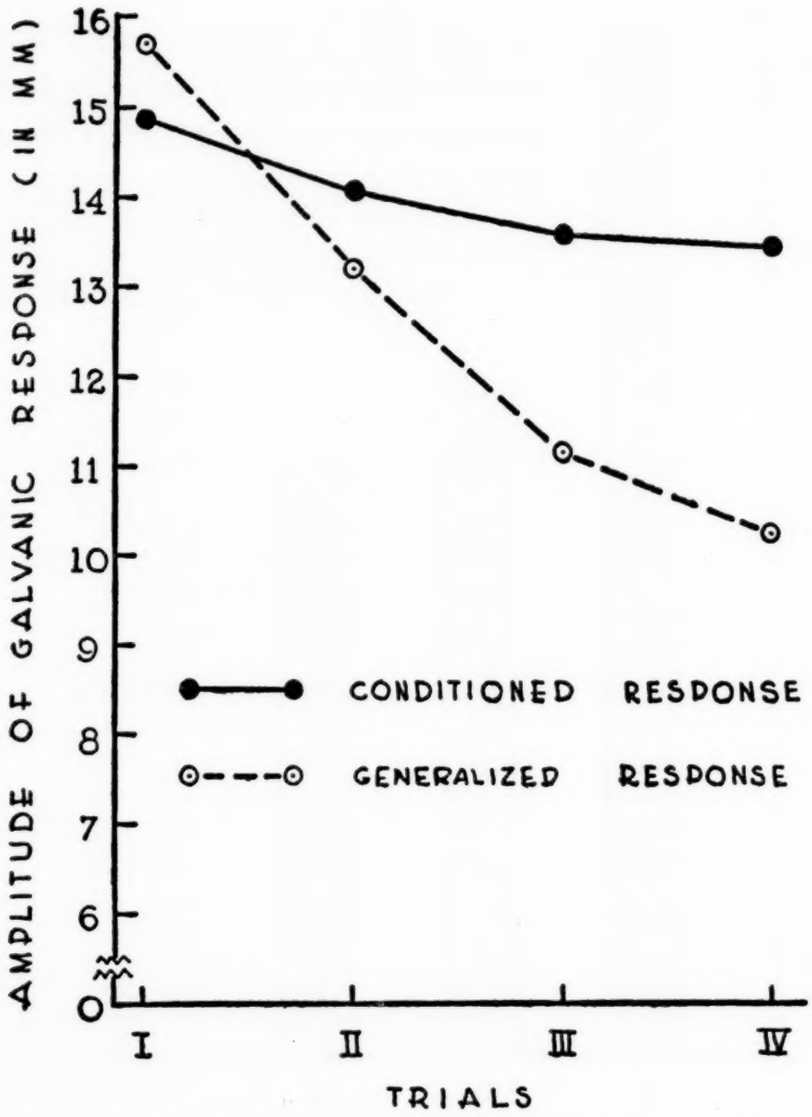

FIG. 2. Magnitudes of conditioned and of generalized galvanic responses on successive test trials following sixteen paired presentations of vibrator and shock.

increase in the amount of generalization. This phenomenon was tested by means of an auditory conditioned response.

'DISINHIBITION' OF CONDITIONED AND OF GENERALIZED RESPONSES FOLLOWING EXTINCTION TESTS

The initial part of this experiment was identical with the procedure described above, except that tonal stimuli were used in conditioning. The sound stimuli were chosen 150 j.n.d.'s apart in intensity. These intensities were 40 and 86 decibels above threshold. The frequency of the tone was

1000 cycles. Forty subjects were employed, half being conditioned to the 40 db tone and half to the 86 db tone.

The experimental procedure is given below:

Step

I. Test response to 40 and 86 db tones six times each.
II. Give 16 reinforcements to either 40 or 86 db tone.
III. Test responses to 40 and 86 db tone four times each.
IV. Suddenly extinguish lights in room in which subject was seated.
V. Test response to 40 and 86 db tones.

The test trials were obtained in the counterbalanced practice order described in the previous experiment.

The conditioned and generalized responses on the first four test trials are given in Table 4.

TABLE 4

MAGNITUDES OF GALVANIC RESPONSE (IN MM.) TO CONDITIONED INTENSITY OF SOUND AND TO ONE 200 J.N.D.'S REMOVED, ON SUCCESSIVE TEST TRIALS FOLLOWING 16 REINFORCEMENTS

Based upon 40 subjects: half conditioned to 86 db tone; half conditioned to 40 db tone.

	Test Trials			
	1	2	3	4
Group Conditioned to 86 db Intensity of Tone				
Conditioned responses	12.61	12.49	11.78	11.91
Generalized responses	10.55	8.19	7.80	7.49
Group Conditioned to 40 db Intensity of Tone				
Conditioned responses	11.59	11.10	9.99	9.71
Generalized responses	16.30	13.10	11.39	8.50
Combined Groups				
Conditioned responses	12.10	11.79	10.88	10.81
Generalized responses	13.42	10.64	9.59	7.99

It will be observed that by the fourth test trial the magnitude of the conditioned response is 10.81 mm, and that of

the generalized response is 7.99 mm. When the lights were suddenly extinguished in the room in which the subject was seated and responses obtained immediately thereafter, the magnitudes of the two responses became:

Conditioned response	13.92 mm
Generalized response	13.79 mm

The effect observed is completely in accord with the theoretical prediction. At the end of the testing before disinhibition the ratio of the conditioned to the generalized responses was 1.35, but after disinhibition it declined to 1.01, indicating marked increase in the amount of generalization. The results here presented concerning disinhibition of the conditioned galvanic skin reactions are in complete agreement with those of Wenger (**9**) and of Switzer (**8**).

The difference in the speed of extinction of the conditioned and of the generalized responses in both the vibratory and auditory experiments is very interesting. These observations support an earlier study of the writer (**4**) in showing that the conditioned response is highly stable but that the magnitude of the generalized response falls very rapidly. On the first test trial the two responses are almost the same size, but by the fourth test trial the ratio of conditioned response to generalized response is 1.35. The rapid extinction of generalized responses is also reported by Liddell, James and Anderson (**6**). The present observations concerning the relative stability of the conditioned and generalized responses seem related to the experiments of Brogden and Culler (**2**). These investigators found that responses based upon the primary motivation were about four times as resistant to extinction as those based upon higher-order motivations.

Difficult to explain is the fact that on the first test trials the responses to the novel stimuli are as large as or larger than those to the conditioned stimulus. Pavlov (**7**) states that on the first test trials with a novel stimulus of the same sensory modality, a very small response is obtained, due to external inhibition. Upon the first presentation of the novel stimulus the investigatory reflex is supposedly elicited, and this response inhibits the established salivary response. If

novel stimuli also produce external inhibition of the galvanic skin response (Wenger, **9**), we should expect the same phenomenon in the present case. Instead, in the present experiment the generalized response is larger than the conditioned. Preliminary studies have shown that the omission of the pre-conditioning tests with the novel stimuli results in an even larger generalized response after conditioning. Further experimentation is indicated.

It is possible that the difference in the nervous mechanisms involved may account for the discrepancy of results. The dual mechanism in the case of the salivary system is well known. If the chorda tympani is electrically stimulated saliva is increased; but if the sympathetic component is stimulated, a diminution in flow results. External inhibition in Pavlov's dogs might thus be explained as the effect of the antagonistic action of the sympathetic component resulting from startling sensory stimuli. On the other hand, startling sensory stimuli typically increase the magnitude of galvanic response.[3]

The difference in method of measuring the response may also be a factor. In Pavlov's experiments the salivary response is measured over a considerable period, *e.g.*, 30 seconds. The galvanic skin reaction, on the other hand, is measured as an immediate response. If novel stimuli have two effects, analogous to 'facilitation' and 'extinction' following cortical stimulation,[4] the immediate response would be augmented, while the response measured over a prolonged period would be diminished. It would then be predicted that if an experiment were conducted on the galvanic skin reaction, in which the external and the conditioned stimuli were simultaneously presented, no external inhibition would be obtained. This outcome would be in accord with the results of Hilgard [5] on reflex responses, which show reinforce-

[3] *Cf.* DARROW, C. W., The palmar galvanic skin reflex (sweating) and parasympathetic activity, *Amer. J. Psychol.*, 1936, **48**, 522–524.

[4] DUSSER DE BARENNE, J. G., AND W. S. MCCULLOCH, An 'extinction' phenomenon on stimulation of the cerebral cortex, *Proc. Soc. Exper. Biol. and Med.*, 1934, **32**, 524–327.

[5] HILGARD, E. R., Reinforcement and inhibition of eyelid reflexes. *J. General Psychol.*, 1933, **8**, 85–113.

ment with simultaneous presentation, but subsequent inhibition.

Interaction of Conditioned and Extinguished Responses on Successive Test Trials

The counterpart of the experiments reported above, concerning the analysis of the changes in the magnitudes of conditioned and of generalized responses, involves the testing of two conditioned responses following the extinction of one of them. The first experiments are related to generalization of conditioned excitatory tendencies, the present section to generalization of extinctive inhibition.

Conditioned responses were established concurrently to two tones, 75 j.n.d.'s apart in frequency, equated for loudness. The first tone was 153 cycles, and the second 1967. The tones were equated to a tone of 1000 cycles with a sensation level of 40 db. Of the twenty subjects in this experiment, ten were conditioned to the 153 cycle tone, and ten to the 1967 cycle frequency. The complete program was as follows:

Step

I. Test response to tones of 153 and 1967 cycles six times each.
II. Give 8 reinforcements to each of the tones in prearranged order.
III. Test responses to each of the two tones.
IV. Give 16 unreinforced presentations of either the 153 or 1967 cycle tone.
V. Test responses to the two tones four times each.

The responses to the two tones before conditioning were 3.68 mm of galvanic deflection for the first group, and 3.37 mm for the other. No reliable difference in the two groups is revealed.

The significant portions of the results are the data presented in Table 5, showing the magnitudes of galvanic response on successive test trials of the response conditioned and then extinguished and of the response conditioned but not extin-

guished. From these data, which are graphed in Fig. 3, one can observe that the responses which were conditioned but not extinguished decline as they would in a normal experi-

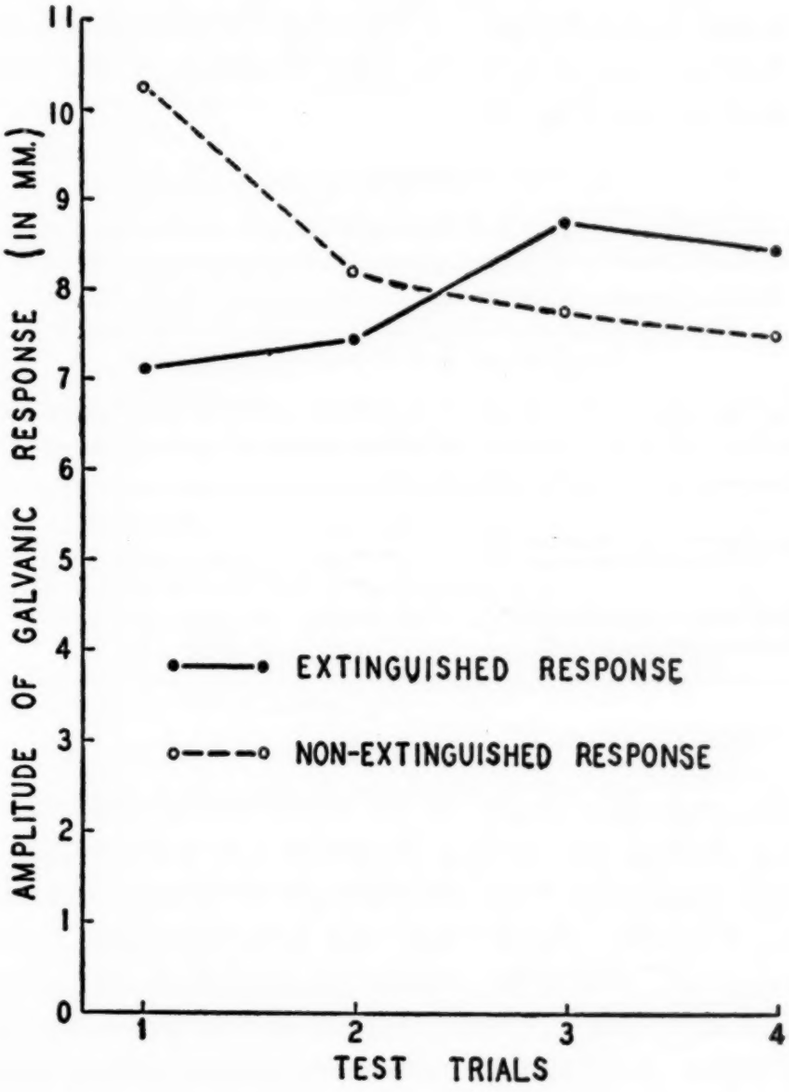

FIG. 3. Magnitude of galvanic response on successive test trials (1) to tone A used as conditioned stimulus and then given repeatedly without reinforcement and (2) to tone B used as conditioned stimulus but not previously presented without reinforcement.

mental extinction program. This is to be expected, since the test trials were unreinforced. But the increase in response to the frequency of tone which was given repeatedly without reinforcement after conditioning is interesting. The most probable explanation is that the extinction which had been

achieved by the unreinforced stimulations was disinhibited by the initial presentations of the frequency which had never been given during the extinction series. This outcome would be predicted from known evidence as to the nature of extinction and disinhibition. Following the disinhibition, which is, as Pavlov has shown (7) very transitory, the responses again decline (see Fig. 3).

TABLE 5

AVERAGE MAGNITUDES OF GALVANIC RESPONSE (IN MM.) ON SUCCESSIVE TEST TRIALS TO FREQUENCY OF STIMULUS WHICH WAS CONDITIONED AND THEN GIVEN REPEATEDLY WITHOUT REINFORCEMENT UNTIL RESPONSE BECAME PARTIALLY EXTINGUISHED, AND TO FREQUENCY CONDITIONED BUT NOT EXTINGUISHED

Half of the subjects in each group were given unreinforced presentations of 1967 cycle tone; half were given unreinforced presentations of 153 cycle tone.

Prior Experimental Procedure for Conditioned Responses	Test Trials			
	1	2	3	4
Conditioned, then extinguished.....	7.09	7.45	8.78	8.52
Conditioned, not extinguished.....	10.26	8.20	7.76	7.55

SUMMARY

1. An analytical study of the conditioned and generalized responses during the testing program was undertaken. Galvanic skin responses were conditioned to tactile and auditory stimuli. Electric shock was the unconditioned stimulus. The temporal relations were: conditioned stimulus, 400 milliseconds; pause, 90 milliseconds; shock, 75 milliseconds.

2. Tactile conditioned responses extinguished comparatively slowly during the test trials following reinforcement. The extent of generalization was determined by testing the response to a vibratory stimulus differing from the conditioned stimulus in intensity. Half of the subjects were conditioned to the strong and half to the weak intensity. Generalized responses to a vibratory stimulus of different intensity declined in magnitude much more rapidly than the conditioned responses during testing.

3. If the more rapid decline in the magnitude of the

generalized response was due to a greater amount of experimental extinction resulting from the testing procedure, more spontaneous recovery should be expected for the generalized than for the conditioned responses. Testing of conditioned and generalized auditory responses after 24 hours confirmed this deduction, since the generalized responses showed more marked spontaneous recovery than did the conditioned responses.

4. Disinhibition of the extinguished generalized responses was demonstrated when the lights in the room were suddenly turned off after some degree of differentiation between the conditioned and generalized responses had been established.

5. The result that on the first test trials the generalized responses were larger than the conditioned responses is in conflict with the finding of Pavlov that the initial generalized responses are very small. It is suggested that the differences in the neural control of salivation and galvanic responses may explain this discrepancy. The effect of novelty, which produces the Pavlovian 'what-is-it' reflex, may arouse a response antagonistic to salivation, but produce positive summation with galvanic responses.

6. When conditioned responses were established to two widely separated frequencies of tone, and subsequently one was extinguished, testing resulted in the extinction of the response not previously extinguished, but an initial augmentation of the extinguished response, presumably as the result of 'disinhibition' by the other test stimulus. With further testing this effect disappears and the extinguished response again begins to decline.

(Manuscript received December 14, 1936)

References

1. Bass, M. J., and C. L. Hull, The irradiation of a tactile conditioned reflex in man, *J. Comp. Psychol.*, 1934, **17**, 47–66.
2. Brogden, W. J., and E. Culler, Experimental extinction of higher-order responses, *Amer. J. Psychol.*, 1935, **47**, 663–669.
3. Hovland, C. I., The generalization of conditioned responses. I. The sensory generalization of conditioned responses with varying frequencies of tone, *J. General Psychol.* (in press).

4. ——, The generalization of conditioned responses. II. The sensory generalization of conditioned responses with varying intensities of tone (in prep.).
5. ——, Magnitude of conditioned galvanic skin responses with varying intensities of conditioned stimulus (in prep.)
6. LIDDELL, H. S., W. T. JAMES, AND O. D. ANDERSON, The comparative physiology of the conditioned motor reflex: based on experiments with the pig, dog, sheep, goat, and rabbit, *Comp. Psychol. Monog.*, 1934, **11**, No. 1, pp. 89.
7. PAVLOV, I. P., *Conditioned reflexes* (Trans. by G. V. Anrep), London: Oxford Univ. Press, 1927, pp. xv + 430.
8. SWITZER, S. A., Disinhibition of the conditioned galvanic skin response, *J. General Psychol.*, 1933, **9**, 77–100.
9. WENGER, M. A., External inhibition and disinhibition produced by duplicate stimuli, *Amer. J. Psychol.*, 1936, **48**, 446–456.

AN ONTOGENETIC STUDY OF CEREBRAL ELECTRICAL POTENTIALS IN THE GUINEA PIG

BY H. H. JASPER, C. S. BRIDGMAN, AND L. CARMICHAEL

The Bradley Home, Brown University, and the University of Rochester

This paper reports an exploratory investigation of changes in the rhythmic electrical potentials of the brain of the guinea pig, as a function of age.

Previous investigations of the character of potentials from immature and developing brains have been few in number. Kornmueller (**6**) mentions that the rabbit gives only slight indication of brain activity, and for the first time, six days after birth. By twelve days, he reports, development is nearly complete in all parts of the cortex.

Other experimenters have investigated the character of brain potentials of children, but as yet, no complete report of such a study has been published (**1**, **7**, **8**).

The amplifiers, oscillographs and continuous paper film recording system used in this study are those previously described by Jasper and Andrews (**5**).

For the most part, the electrodes used were fine silver wires with small spherical enlargements at the contact end. These electrodes were insulated with a coating of celloidin. The electrodes were inserted through small keyhole-shaped openings in the skull to make contact with the underlying dural membrane. They were then held in place by small celluloid wedges.

In general, 'bipolar' electrode recording was used, although in some cases an additional 'monopolar' record was taken, employing a diffuse electrode, located on the ear. The paired electrodes were placed 5–10 mm apart in the parieto-occipital region, about 7 mm to the left of the midline. These measurements varied, depending on the size of the animal, being somewhat greater for the larger than for the smaller animals.

The guinea pigs were healthy members of the general laboratory colony. When using prenatal animals, age was calculated for the most part from records of the date of insemination. In a number of cases where this information was not available, age was estimated from Draper's tables of length, weight, and age relationships for fetal guinea pigs (**3**). These prenatal animals were examined while umbilical circulation remained intact, although in some instances, air-breathing was established in late fetuses. The young postnatal guinea pigs were born in the colony, their ages thus being available.

In presenting the results obtained from this series of animals, save when discussing the onset of activity, we have used the method of measurement whereby frequency of variation of potential is determined in those parts of the record exhibiting sufficient regularity for the measurement of such variations. We have ordinarily considered the criterion of sufficient regularity to be fulfilled only in those parts of the record showing at least five successive cycles whose durations do not differ from each other by more than approximately 10 percent. In applying this criterion, the various durations to be included in each frequency measurement were compared by using dividers, it being inaccurate to rely on unaided visual estimation.

Of the 11 litters of fetuses examined, those younger than 48 days gestation age, consisting of 11 fetuses from five litters 35, 38, 43, 44, and 45 days old, exhibited no brain potentials. Fig. 1 is a sample of the record taken from one of the 38 day animals. In this and the following figures, the upper portion of each record indicates time in 20ths of a second. The horizontal lines placed between the various records indicate durations of one second. The figures opposite the vertical lines at the ends of the records indicate the number of microvolts required to produce a deflection equal to the length of the line. In sample '*A*,' the first line represents the record from paired electrodes on the cortex, and the second line was taken from one of these electrodes to a diffuse lead on the nose.

The difficulties involved in firmly placing electrodes in

these young animals, whose skulls are, at this stage of development, still not fully ossified, added to the limitations imposed by the necessity of maintaining maternal circulation, have made it impossible to secure records entirely free from irregularity introduced by artifacts. However, these disturbances may be clearly distinguished from the brain potentials of more mature animals.

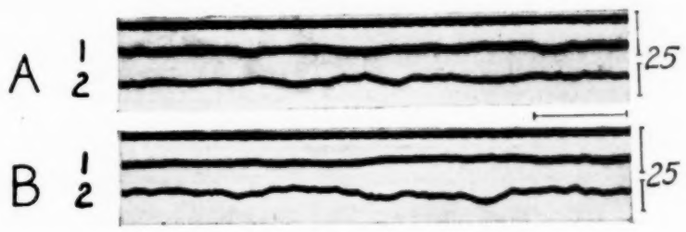

FIG. 1.

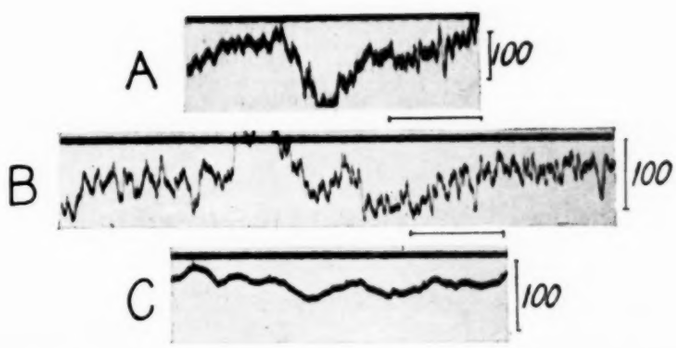

FIG. 2.

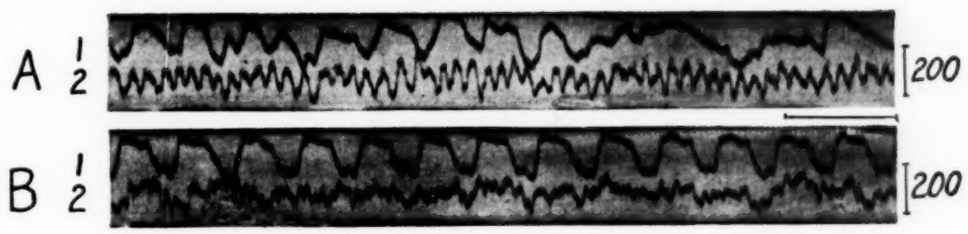

FIG. 3.

At the conclusion of each period of recording from a fetus with maintained maternal circulation, the umbilical blood vessels were clamped. The resultant of this action presented an admirable experimental control, for soon after the clamp was applied, the bioelectric phenomena of the cortex were always abolished, but artifacts remained unaffected. In the older fetuses, this procedure resulted very shortly in complete

disappearance of the electrical variations being recorded, even before the reflex activity of the animal was reduced, indeed when at times it seemed temporarily augmented. In the relatively early fetuses under discussion, interference with the umbilical circulation, even when continued until all reflexes were abolished, resulted in no alteration of the character of the records. That such base-line deviations as were found in the records of fetuses less than 48 days gestation age persist unchanged and undiminished after the animal has reached a state where normal brain activity is no longer present is excellent proof that these incidental variations did not at any time originate from brain activity. Thus, in Fig. 1, sample '*B*' is taken after clamping the umbilical cord, and shows the same minor deviations as that taken from the normal animal.

The next litter in age, consisting of three fetuses at 48 days, gave some records containing the earliest evidence of brain activity recorded in this series. Two members of the litter gave negative records similar to those recorded from younger animals. From the first fetus to be examined in this litter, however, there was obtained a record which, for short intervals, from time to time exhibited potentials similar in certain characteristics to the brain potentials of the adult guinea pig. In Fig. 2, record '*A*' shows a typical sample of this activity. In some instances, measurable frequencies were present, ranging from 7–20 waves per second.

Of two fetuses 54 days of age, one gave a record very similar to that just described. Its litter mate, on the other hand, presented a record of continuous potential waves characteristic of the adult. Regular rhythms were present, ranging in frequency from 5 to 34 cycles per second. In addition, more rapid but irregular waves could be distinguished, and the record is marked throughout by large amplitude, long duration waves not regular enough to be treated by the procedure outlined above. In Fig. 2, sample '*B*' is typical of this record. Sample '*C*' is taken about 50 seconds after shutting off umbilical circulation. Disappearance of the potentials as a result of this operation, at a time

when conditions making for artifactual variations remain unchanged, indicates that they depend on normal brain circulatory conditions.

Three fetuses of a litter which had reached the age of 56 days showed records of continuous brain potentials.

From the above presentation of results, it is evident that continuous brain potentials first appear in the guinea pig in fetuses as young as 54–56 days, with positive indication of activity as early as 48 days. This period is about two weeks before normal birth time.

The data on the developmental course of brain potentials is presented in Table 1, which includes a summary of the

TABLE 1

	Most Characteristic Frequencies		2nd Most Characteristic Frequencies		Total Frequency Range
	Range	Amplitude	Range	Amplitude	
Age in Days of Gestation					
48	7–9	19	—	—	7–20
54	17–23	18	—	—	5–34
60	11–14	30	—	—	5–20
61	15–18	25	—	—	9–28
64	9–18	27	—	—	6–41
65	11–20	18	21–30	15	8–50
Post-natal Age					
3 days	16–19	50	32–39	35	6–39
6 days	15–19	16	—	—	6–53
7 days	13–16	40	—	—	13–33
11 days	13–16	60	—	—	3–20
1 month	10–11	50	14–18	24	5–30
2 months	20–30	20	—	—	8–50
3.5 months	7–8	75	36–40	30	7–48
5 months	7–8	35	31–36	20	7–39
8 months	13–21	40	32–33	30	6–40
Adult	8–10	225	31–40	50	8–40
Adult	—	—	—	—	6–35
Adult	7–9	60	41–46	18	6–47
Adult	10–17	70	35–41	30	7–41

measurements of frequencies found at the various age levels examined, including those from four adult animals. Under the heading, 'most characteristic frequencies,' is presented the range within which the greatest grouping of frequency measurements occurred, at each age level, along with the average amplitude in microvolts of the waves within this

range. Under the heading, 'second most characteristic frequencies,' is presented the range within which any secondary grouping of measurements occurred, along with their average amplitude. These ranges were determined by the examination of complete tables made up of the various frequency measurements at each age level. The last column in the summarizing table presents the range which includes all the frequencies at each age studied. Examination of this table reveals no definite ontogenetic trend either in the limits of the total frequency range, or in the frequencies which are most characteristic of the different ages. There is, however, an indication that a secondary grouping appears more regularly in the older animals.

In addition, certain qualitative age differences with respect to regular frequency may be described. Long bursts of regular waves are found only in the older animals, and the measurable portions of the records are greater in number than in younger animals. Further, in the young animals, those parts of the record which exhibit a measurable frequency fit less readily within the criterion of regularity previously described. That is, the rhythmic portions of the records from mature animals are more perfectly regular, exhibiting durations more nearly exactly comparable, and showing less interference from other frequencies.

A further differentiation between the very earliest records and those from older animals is that, as noted above, at the very onset of activity, the brain potentials are present discontinuously.

Reference to Table 1 will bring out the fact that there is, in general, an increase in amplitude of the most characteristic frequencies as age increases. This finding is subject to the usual limitations of accuracy present in the measurement of amplitude, arising from the impossibility of insuring exactly comparable shunting effects from one electrode placement to another, but the trend seems sufficiently marked to justify notice.

While recording from each animal, various sound and light stimuli were presented, the experiment being carried on in an

otherwise dark or dimly lighted sound-resistant room. In numerous instances, definite changes in the character of the records are found, directly following stimulation. Although these changes are not invariably found, the contention that they are brought about as a result of the stimulation is strengthened by the fact that similar changes have recently been reported in the cat (4). These effects of stimulation involve either a change in frequency, a change in amplitude, or a complete change in the pattern of electrical response from that exhibited prior to stimulation. Definite changes in the character of brain potentials following stimulation have been found in fetuses as early as the 60th day of gestation.

Certain of the records from older guinea pigs may be taken as illustrative of the effects of general excitatory conditions. In recording from these animals, a restraining holder was first employed, in which situation the animals struggled almost constantly. Records under these conditions exhibited relatively greater numbers of higher frequencies, with almost no slow regular potentials. After release of the animal, the record gradually developed more of the slower potentials usually present in the cortical electrogram of the normal unrestrained adult guinea pig.

In each of the records of this study, there are some waves comparable to the characteristic human brain activity called the alpha rhythm. In most cases, this 8–12 per second frequency is not outstandingly characteristic of the record as a whole. It should be pointed out, however, that it is possible to record from the guinea pig brain, electrical variations which in frequency, regularity, and continuity, present the same nature as a good record of alpha rhythm from the human cortex. The adult guinea pig of whose cortical electrogram Fig. 3 is an example, presented a record marked throughout by long periods of activity of the sort illustrated by *A*-2. *B*-2 represents the nature of the record during the intervals when the slow regular activity was not present. (The large slow deviations opposite Fig. 1 in both *A* and *B* are artifactual in nature, being recorded from a frontal region which was observed to be pulsating at the same rate.) The brain

potentials recorded in the second line of both *A* and *B* are characteristic of the adult guinea pig. In other adult records, however, the alpha rhythm was not present with the same preponderance.

Discussion

It has been one of the aims of this investigation to correlate the development of brain activity with other maturational indices. Lack of information concerning the histological development of the guinea pig's brain makes it impossible at present to correlate the onset of electrical activity with any specific changes in the brain. However, it may be pointed out that during or shortly before the period of initiation of brain potentials found in this study, that is, 48–56 days, certain behavioral indications of an assumption of function by higher parts of the nervous system may be observed. One of the collaborators in this study has previously reported finding a marked increase in responses after high transection of the spinal cord in fetuses as early as the 44th day (**2**). The motor mechanisms of the eye are well developed at the period under discussion, and the eyes are first open naturally at this time. In addition, the period of onset of electrical activity precedes by only a few days at the most the striking reduction of stimulus-released activity characteristic of late stages of fetal life.

The findings of well-developed alpha rhythms in the guinea pig would lead one to believe that this phenomenon is connected with some basic neurological mechanism, rather than with any higher elaboration of nervous function found only in the primates.

Conclusions

1. The brain potentials of the guinea pig first appear when the age of 48–56 days of gestation has been attained.

2. No quantitative ontogenetic trend in the characteristic frequency of potential variation, nor in the total range of frequencies has been discovered. The appearance of secondary groupings of characteristic frequencies is found more often as age increases.

3. The average amplitude of the characteristic frequencies at any age shows an irregular yet definite increase with age.

4. Definite though not invariable effects of stimulation on the character of the cortical electrogram have been noted, occurring as early as the 60th day of gestation.

5. The guinea pig brain first exhibits electrical activity at a time when behavioral indications also point to maturation of higher nervous centers.

(Manuscript received December 2, 1936)

References

1. Berger, H., Uber das Elektrenkephalogram des Menschen. V, *Arch. f. Psychiat., u. Nerv.*, 1933, **98**, 231–254.
2. Carmichael, L., An experimental study in the prenatal guinea pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life, *Genet. Psychol. Monog.*, 1934, **16**, 5–6, 337–491.
3. Draper, R. L., Prenatal growth of the guinea pig, *Anat. Rec.*, 1920, **18**, 369–392.
4. Jasper, H. H., Cortical excitatory state and synchronism in the control of bioelectric autonomous rhythms, *Symp. Quant. Biol.*, Cold Spring Harbor, 1936, **4**, 320–338.
5. Jasper, H. H., and Andrews, H. L., Human brain rhythms. I. Recording techniques and preliminary results, *J. Gen. Psychol.*, 1936, **14**, 98–126.
6. Kornmueller, A. E., Die bioelectrischen erscheinungen architektonischer felder der grosshirnrinde, *Biol. Rev.*, 1935, **10**, 383–426.
7. Lindsley, D. B., Brain potentials in children and adults, *Science*, 1936, **84**, 354.
8. Loomis, A. L., Harvey, E. N., and Hobart, G., Electrical potentials of the human brain, *J. Exper. Psychol.*, 1936, **19**, 249–279.

CARDIAC ACCELERATION IN EMOTIONAL SITUATIONS

BY J. G. BEEBE-CENTER AND S. S. STEVENS

Harvard University

This study presents an attempt to extend our knowledge of autonomic factors in emotion by studying the cardiac behavior of cats in emotional situations. The specific experimental problem was to ascertain precisely the temporal course of the heart rate—how long after stimulation before the rate begins to increase (latent time), how long before the maximum rate is reached (action time), and how long before the rate returns to normal (recovery time).

The apparatus consists essentially of copper electrodes attached to either side of the cat's thorax by means of a harness, a voltage amplifier, a power amplifier, and a Western Union undulator. The undulator records on ticker tape the amplified action potential of the heart and is equipped with an additional marker to write time signals (seconds) and signals indicating the onset of stimulation. Except in a few cases, which will be specially noted, all records were taken with the cat in a wire box 1.8 ft. × 0.8 ft. × 1.3 ft.

Figure 1 gives a sample record obtained upon firing a shot.

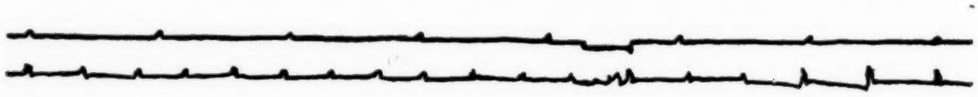

FIG. 1. Sample of record. Cat *F*. Stimulation with shot. Upper line: time in seconds and signal corresponding to stimulation. Lower line: heart beats. Read from right to left.

The procedure in deriving temporal data from such a record was primarily to measure the distance between the excursions corresponding to heart beats by means of a millimeter scale, and also to measure the distance between second marks. (The latter measurement was necessary because the speed of the tape on the recorder was varied to suit experimental

requirements, and further, was not strictly constant for a single setting of the apparatus.) Distances so measured were translated into beats per minute. The temporal position of a beat is always given by the end of the beat. This procedure lengthens all temporal measures somewhat (1/10 to 1/4 of a second, depending upon the rate) relative to what they would be if measurement were made to the center of the beat. Since the most striking feature of the latent and action times is their shortness, this conservative procedure is the more desirable.

The Reaction in General

Figures 2 and 3 are typical graphs representing individual rates of heart beat in the case of a shot stimulus, which

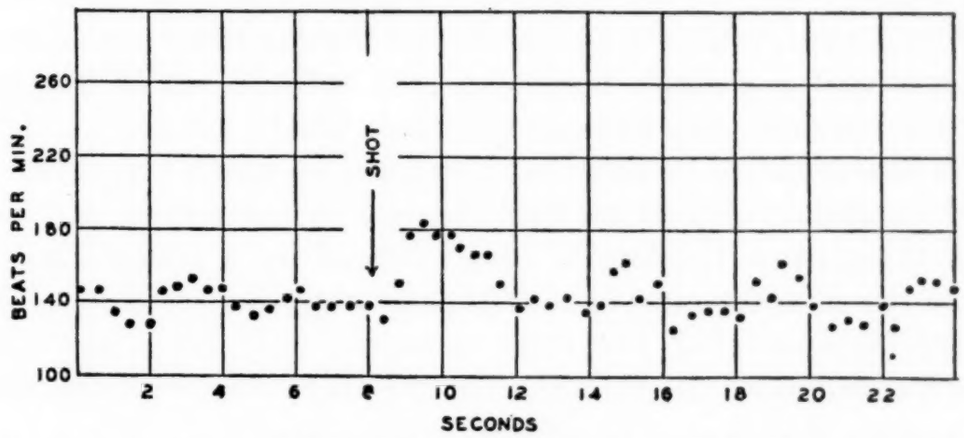

Fig. 2. Temporal course of the cardiac response to shot. Each point represents the rate of beating as determined from the time between two successive beats.

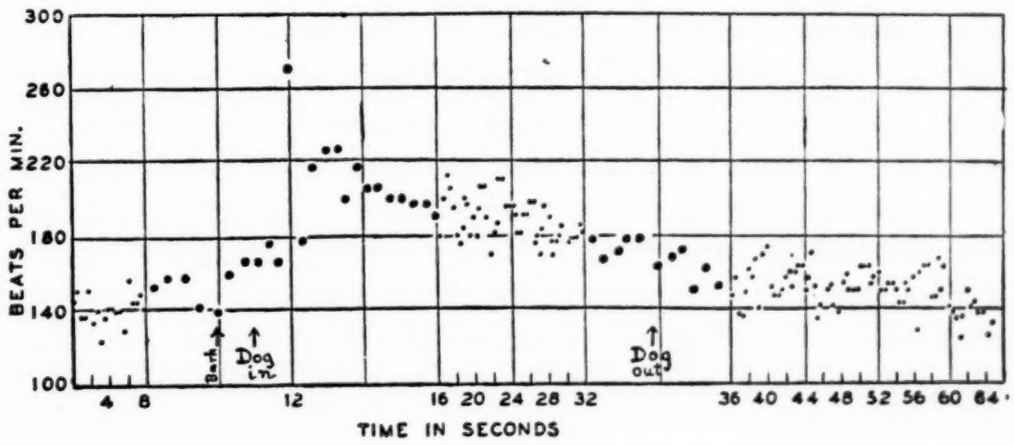

Fig. 3. Temporal course of cardiac response to dog. Each point represents the rate of beating as determined from the time between two successive beats. Note that the abscissa scale is not uniform, but is increased fourfold at the beginning and end of the reaction in order to bring out details.

produces a minor rise (Fig. 2), and in the case of presentation of a dog, which produces a fairly marked rise (Fig. 3). Fig. 2 shows a brief latency (about .5 sec.), a rise, rapid at first and then slower, a maximum 1.55 sec. after stimulation, and then a gradual return to normal over a period of about 2.5 seconds. Fig. 3 represents a situation not wholly comparable, for here the stimulus (dog) is not brief, but persists for 23 seconds. Insofar, however, as comparison is legitimate, there is considerable similarity. Onset of stimulation (the bark of the dog) produces, after a latency under .5 sec., a rapid rise in heart rate, tapering off to a maximum about 3 seconds after stimulation; and the return to normal, which begins long before cessation of the stimulus, is again gradual. One feature, however, distinguishes this graph from that for a shot. About 1 sec. after the appearance of the dog there is a 'spike,' representing a single beat at a rate entirely out of keeping with previous and subsequent beats (extra systole?). This feature occurred in three other records in which presentation of an animal elicited marked changes in heart rate. In two of these cases the 'spike' was followed by a resumption of the rise in heart rate, in the other by a gradual decline. The circumstance that the spike occurred only upon the first presentation of the stimulus on the day of experimentation suggests that it may be related to surprise.

Figures 2 and 3 are typical records. They illustrate the order of magnitude of the temporal features of the cardiac responses with which we worked. The specific problems to be dealt with concern the action time, latency, recovery time and magnitude of the cardiac response. A considerable amount of data is available from reactions to shots, but technical difficulties have reduced the data available from reactions to other animals. When another animal is used as a stimulus it is extremely difficult to determine reliably and accurately the moment of onset of the stimulus. Another difficulty is that movement on the part of the cat frequently mars the record.

Action Time

As shown in Fig. 2, after an appropriate stimulus the heart rate rises to a certain value and then declines. The time between stimulation and the maximum of this rise is the *action time*. Ideally the maximum should be determined by a smooth curve fitted to the points representing individual heart beats, but since such a curve is not easy to draw in all cases, the beat representing the fastest rate (following soonest upon the preceding beat) was used to define the maximum. Because of the form of the rise in heart rate (see Fig. 10), chance variations tend to lengthen rather than shorten the action time as defined by the procedure.

Figure 4 represents for two cats a large number of action

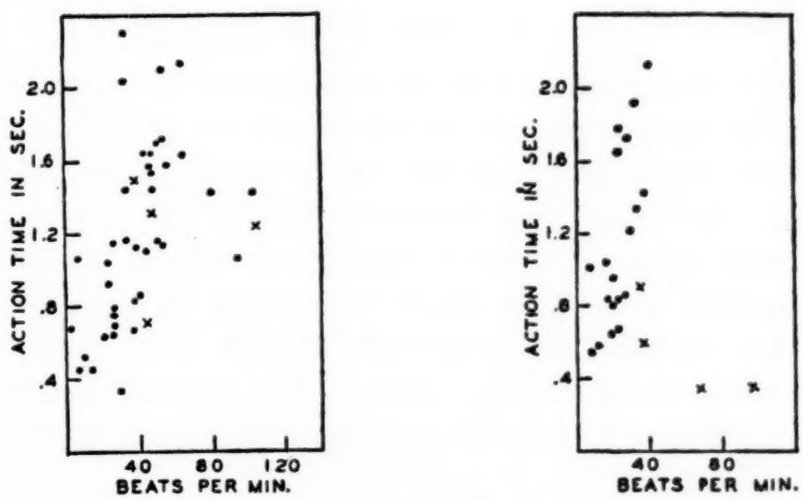

Fig. 4. Relation between action times, *i.e.* interval between stimulation and peak of heart rate (ordinate), and total increment in heart rate at peak (abscissa). Dots represent responses to a shot, crosses responses to stimulation by strange cats or dogs.

Left: Data from cat *A*.
Right: Data from cat *B*.

times to shots and a few to stimulation with cats and dogs. The action times for cat *A* vary from .31 sec to 2.30 sec; for cat *B* from .35 sec to 2.15 sec. For rises of the same general order of magnitude it is clear that, except for a few scattered cases, there is a tendency for action time to vary directly with the amount of change in heart rate. This fact would suggest that rate of increase is fixed and that additional rise in heart rate is effected by increase duration of the rise.

In the case of the shot stimuli, however, the three highest rises in heart rate for cat *A* refuse to fit this picture. With stimulation by cats and dogs, we find that 6 out of 8 cases are exceptions—the rate of rise of the heart beat is considerably greater. We conclude, therefore, that rate of increase of the heart beat is not constant when we compare large and small rises, but rather that it may be increased with increase in the amount of rise involved in the response. Concretely, the rate of rise in the heart beat varies from around 30 or 40 (usual response to shots) to as much as 80 in a second, or even, in the case of 'spikes,' to 300 in a second. (It must be remembered that these 'rates' are computed from only one or two heart beats.)

Latent Time

Latent time is more difficult to measure than action time, because in the responses where onset of stimulus is best controlled—responses to shots—there occurs a skeletal startle response (similar to that demonstrated by Hunt and Landis for human beings)[1] within about .12 sec of stimulation. This response frequently mars the record of the first heart beat after stimulation, especially if the beat occurs very shortly after the stimulus. Hence the data on latent time are more reliable as regards the upper limit of latency than the lower limit.

The table below represents the average, for nine shots, of the times between stimulation and the first beat after stimulation which is faster than the last beat before stimulation. The table also gives the range of these times.

Cat	Avg.	Range
A	.58	.35–1.05
B	.50	.05– .90

The data indicate a latency of about .5 sec. In spite of the shortcomings of the technique, consideration of the many incomplete records which do not figure in the table above suggests that the probable error of this average is less than ± 0.15 sec.

Latent time is defined here as the interval between stimulation and the beginning of the subsequent *rise* in heart rate. Close inspection of the data represented in Fig. 5 indicates, however, that a rise in heart rate may not be the first response to a shot stimulus. The first response in certain

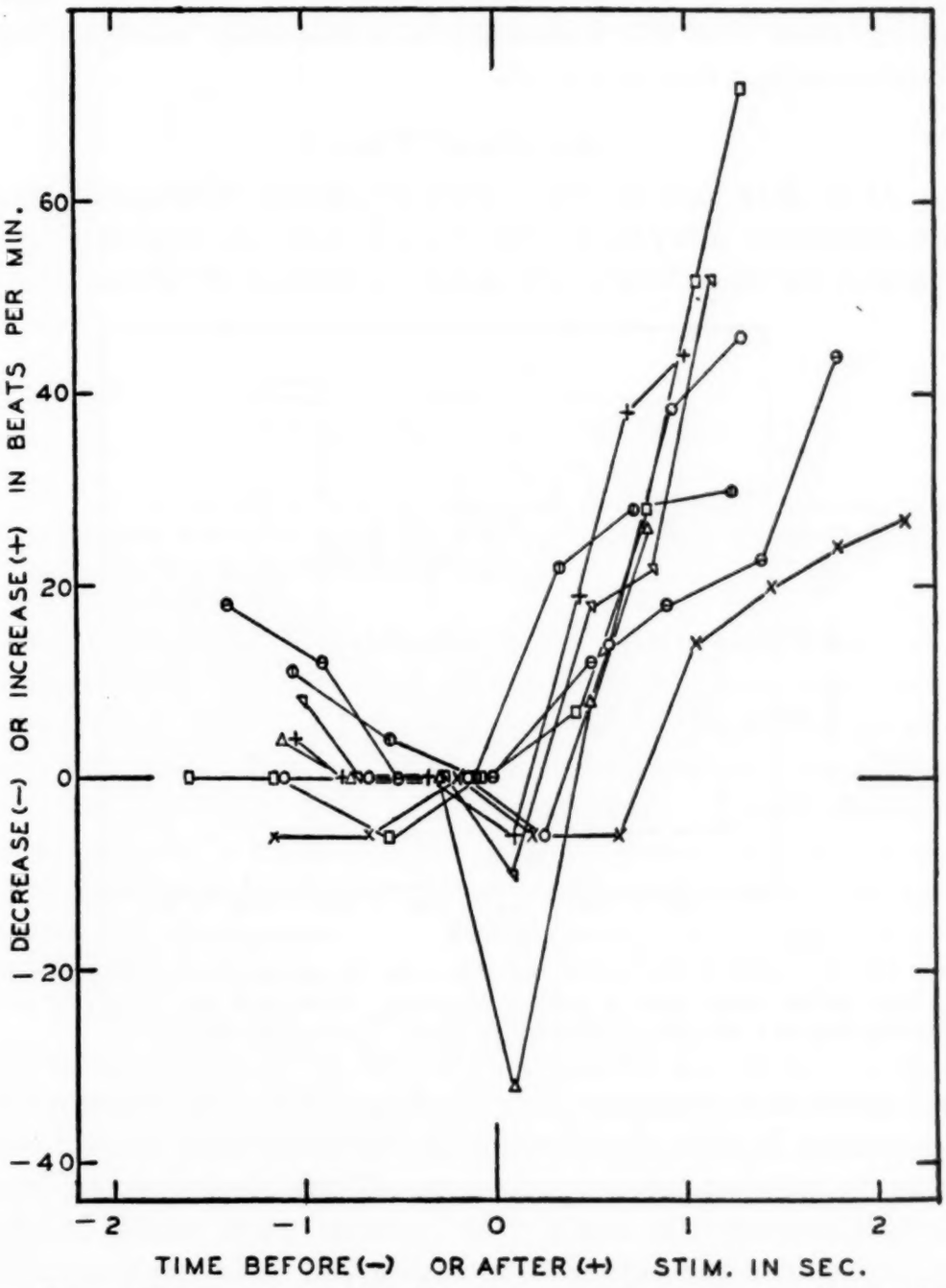

FIG. 5. Temporal course of cardiac responses to shots. Cat *A*. Abscissa: time in seconds before (−) and after (+) stimulation. Ordinate: decrease (−) or increase (+) in rate of beats relative to rate of beat just prior to stimulation.

cases is perhaps a momentary drop, followed by a rise. The occurrence of a drop in rate may depend upon the position in the series of heart beats at which stimulation occurs. If this tentative description can be verified in future experiments we shall have to redefine latent time to take account of the rapid initial drop. Latent time should refer to the interval between stimulation and the beginning of a response, whether that beginning is a rise or a drop.

Recovery Time

It is clear that recovery from emotional stimulation may be extremely delayed, so much so, indeed, as to give rise to various levels of heart rate quite as distinct as those known

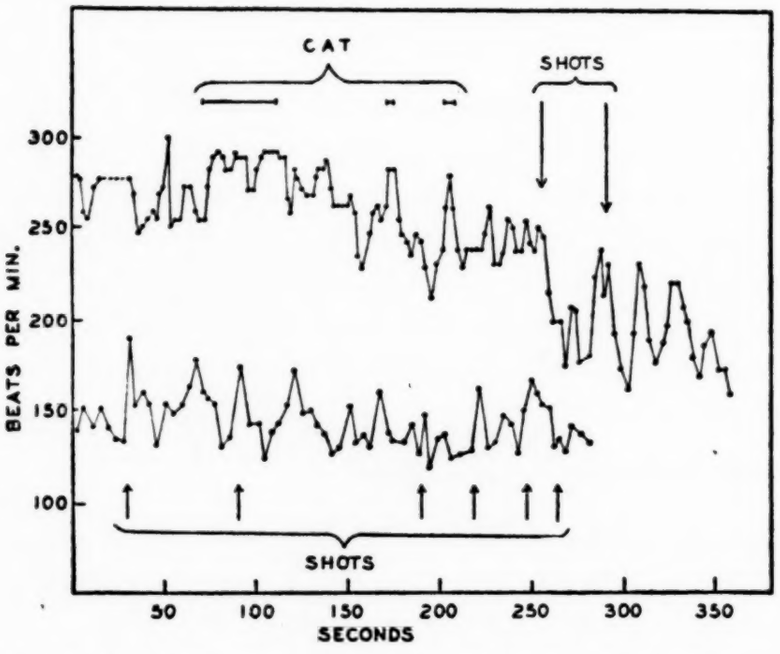

Fig. 6. Levels of heart-rate. Average rates for groups of 10 beats (ordinate) plotted against mean time of groups (abscissa). Data from cat *A*. Upper line: presentations of a strange cat, followed by shots. Lower line: shots.

to occur in connection with feeding. This circumstance is illustrated in Figs. 6 and 7, which show the heart rate of two cats in each of two experiments. Time since feeding was approximately the same (the records were made in the morning, and the cats were always given their single meal in the early evening). The records represent average rate for ten beats plotted against mean time of the ten-beat group.

(Some were taken in connection with a problem not touched on in this paper, namely that of the relation between magnitude of change in heart rate and the nature of the stimulus.)

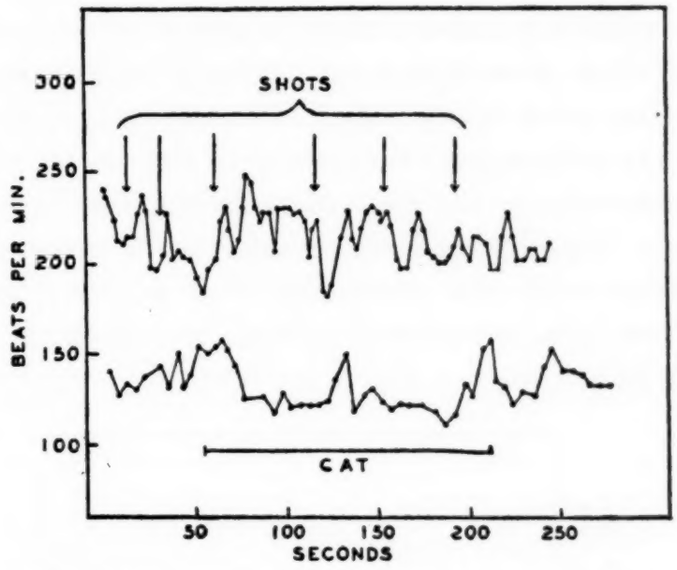

FIG. 7. Levels of heart-rate. Average rates for groups of 10 beats (ordinate) plotted against mean time for groups (abscissa). Data from cat *B*. Upper line: shots. Lower line: presentation of a strange cat.

They show that under different conditions of stimulation the general level of heart rate of the same cat can vary from around 140 per min to around 220 (cat *B*) and from around 150 to around 260 (cat *A*). This means that a base level can be raised by well over 50 percent. (These different levels suggest a number of interesting questions, such as the effect of adrenin in the blood, but in this paper these questions are being disregarded, for the emphasis is on momentary changes.)

In Figs. 6 and 7 two features are of special interest. Although in the lower line of Fig. 6 shots are never followed by marked drops in the level of heart-rate, in the upper line each of the shots is followed by significant drops. (This has nothing to do with the brief startle reaction previously dealt with. The startle reaction is masked by the averaging.) Although we have not sufficient data to discuss these drops advantageously, it appears that one type of stimulation (shots) is able to break up the after-response to another type

of stimulation (cat). The second peculiarity is that whereas presentation of a strange cat to cat *A* results in a rise of heart rate, the same presentation to cat *B* results in a drop. In both cases the stimulus cat spit and showed anger. Cat *B* adopted a wholly passive attitude although he occasionally spit back. Cat *A* adopted an attitude of tension. These facts, together with many other observations, suggest that a rise in heart rate correlates directly with the degree of muscular response, not only in the case of exertion, but in the case of all emotion. This hypothesis remains to be tested.

In contrast with the cases discussed above, the recovery time for most brief emotional stimuli is a matter of seconds rather than of minutes. Fig. 8 represents recovery times for

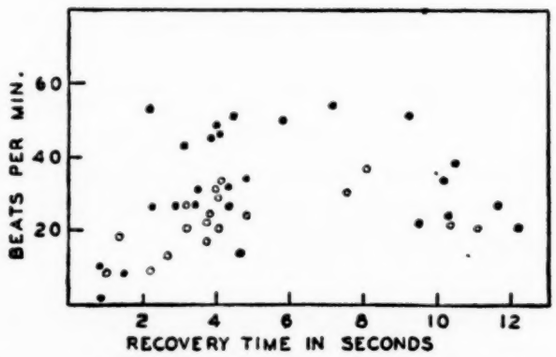

FIG. 8. Cardiac recovery from shot. Magnitude of total increment in heart rate (ordinate) plotted against recovery time, *i.e.* interval between stimulation and first beat beyond peak which is equal to or slower than last beat before stimulation (abscissa). Dots: data from cat *A*. Circles: data from cat *B*. There were three instances only in which the recovery time was greater than 13 seconds.

shots plotted against amount of rise in heart rate in the case of cats stimulated by a shot. By recovery time is here meant the interval between the shot and the first beat beyond the peak whose rate is either equal to or slower than that of the last beat before stimulation. It is apparent that recovery time from shots is usually around 4 sec, since 50 percent of these times lie between 3 and 5 sec. In only three of the cases considered did the time exceed 13 seconds. There seems to be a slight relation between height of rise and recovery time, but it is by no means clear that this represents more than the artifact that high rises in heart rate will tend to occur more frequently when the stimulus falls in a trough

of the normal heart rate variation, and that in this case recovery time will also tend to be longer.

Magnitude of Reaction

Except for the 'spikes' none of the reactions dealt with above is particularly large, as cardiac reactions go in cats. This is perhaps due to habituation on the part of the cats and to the restriction on movement imposed by the apparatus —a restriction which militates against large rises in heart rate.

To illustrate the extreme cardiac reactions of which the cat is capable we may consider data secured with a stethoscope. Fig. 9 shows the heart rate of cat *A* on the first

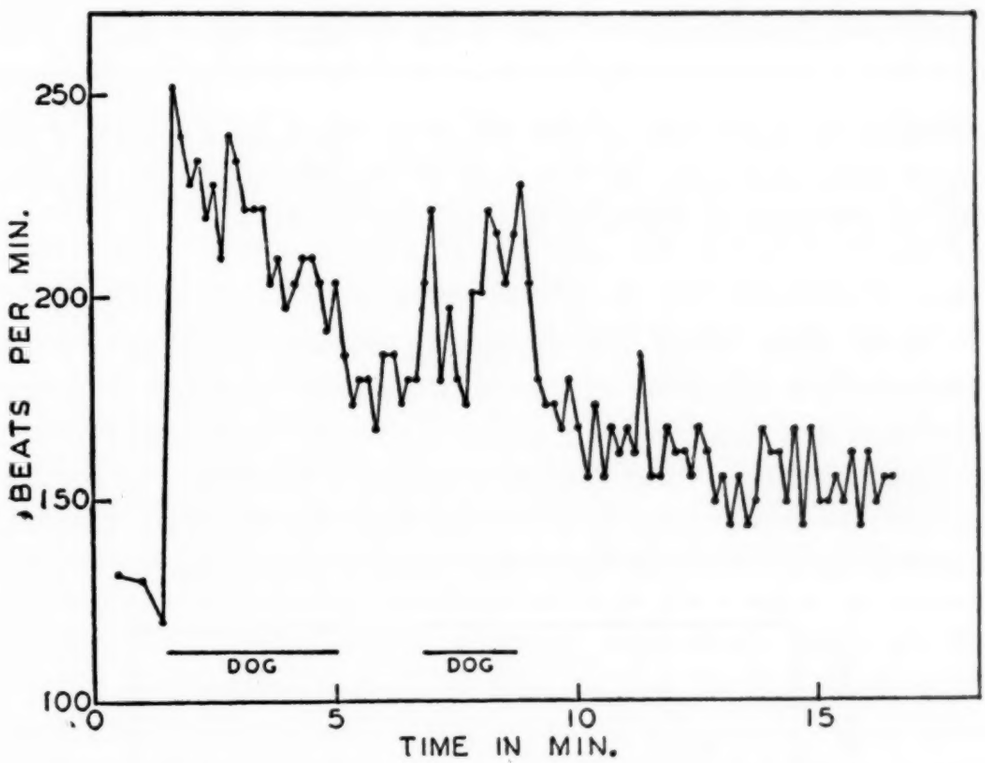

Fig. 9. Cardiac response to dog. First time cat has ever seen dog (Belgian Shepherd). Ordinate: average rate for periods of ten seconds (before stimulation periods are of 30 seconds) recorded by means of a stethoscope. Abscissa: temporal position of periods. Data from cat *A*.

occasion of being shown a large active police dog. The average rise in heart rate for the first two presentations of this dog to cats *A* and *B* was 81 percent of the basal rate.

Three records taken electrically on other cats not previ-

ously used in the experiment, although largely marred by movement, confirm the order of magnitude found by stethoscopic technique. The data from these records are indicated below (slowest beat before stimulation compared to fastest beat after stimulation; slowest group of 10 beats before stimulation compared to fastest group of 10 beats after stimulation). Clearly, then, the magnitude of emotional

TABLE

Cat	Stimulus	Single Beats			Groups of 10 Beats		
		From	to	Diff.	From	to	Diff.
C	Cat *A*	128	246	92%	132	240	82%
D	Belgian Shepherd	170	288	69%	184	284	54%
E	Belgian Shepherd	210	344	64%	212	314	48%

changes in heart rate in the cat may reach 80 percent of the basal rate, not only in the case of individual beats (spikes) but in the case of trains of at least 10 beats.

Discussion

It is clear from the foregoing exposition that we are warranted in speaking of a typical cardiac response to a shot stimulus, and that we can define its major temporal features quite accurately. The typical reaction is presented in Fig. 10. It must be remembered, however, that the diagram involves a certain amount of extrapolation which may require eventual

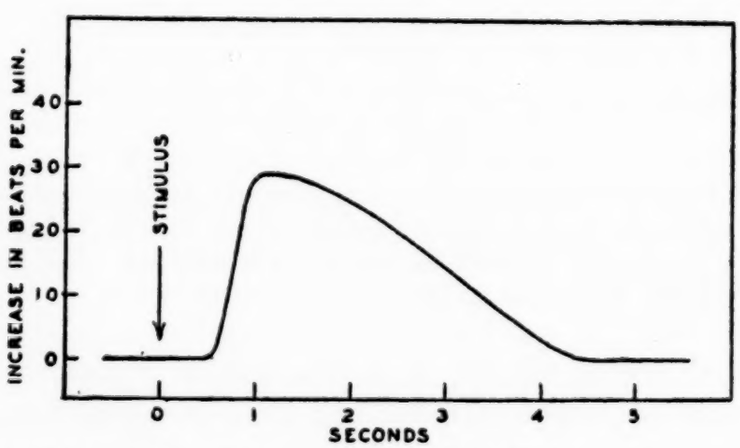

Fig. 10. Schema of cardiac response to a shot. This curve was drawn to illustrate the features of the cardiac response which occur most commonly. The ordinate represents increase in heart rate.

correction. In particular it may be that the latent time will have to be shortened to take account of an initial and extremely brief drop in heart rate, and that there will be found a regular secondary rise or fall in heart rate beyond the initial recovery time. There is in our data some evidence for all of these additional features, but the evidence is not uniform.

The order of magnitude of the rise in heart rate with a shot stimulus is slight compared to the extreme rises which can be elicited by stimulation with cats or dogs. In the latter case the rise may come close to the physiological 'ceiling' for the animal. In the case of such marked rises our data are inadequate for the specification of the temporal features of the response. They do show, however, that the latency and action time are of the same order of magnitude as in the case of the shot response, and in particular, that the greater rise is not due to the same acceleration persisting longer, but to a much greater acceleration.

These data bear upon two related issues. Hunt, in 1899, found that in the anæsthetized dog the latency of cardiac acceleration elicited by stimulation of the saphenous nerve was very brief—apparently less than 1 beat.[2] He found the latency of acceleration elicited by cutting the vagi to be of similar brevity, whereas the latency of acceleration induced by direct stimulation of the accelerators was longer. He further noted that the curves showing acceleration were steep during the early part of stimulation, both in reflex stimulation of the saphenous nerve and in section of the vagus (50 percent of the effect within about 3 beats), but much more gradual during the initial period of stimulation in the case of the accelerators (less than 50 percent of the effect until after about 8 beats following stimulation). With respect to recovery he noted that the stimulation of an accelerator nerve gave a long continuation of the effect after cessation of the stimulus (for 25 beats or so), while acceleration through stimulation of the saphenous nerve did not persist more than some 7 or 8 beats beyond stimulation. These facts support the view, he believes, that in reflex acceleration of the heart, diminution of the tonic activity of the vagi plays the chief role.

In view of the short latencies, action times and recovery times noted by us in the cardiac response to a shot and in the 'spike' reaction to other animals, the question arises whether Hunt's conclusions with respect to reflex acceleration by the saphenous nerve does not hold likewise for these cases of emotional acceleration. Hunt's data are not such as to offer a definite answer to this question.

Fortunately, Rosenblueth and Simeone [3] have recently presented more ample data on the temporal course of vagal and accelerator stimulation in the cat. One of their curves represents the temporal course of accelerator stimulation with a maximal stimulus applied at a frequency of 8 per second. Although this frequency is not strictly the most effective, increase beyond it seemed only slightly more effective. With this form of stimulation the heart eventually reaches a maximum rate of 175 percent of the basal rate. This maximum, however, is reached only after about 20 seconds. After 1 sec the percent of increase is only about 12, after 2 sec only about 22. Now, Fig. 4 shows that for shots, an action time of 1 sec corresponds in cat *A* to an increase in rate of about 35, or, considering his basal rate to be 140, to an increase of 25 percent; for cat *B* to an increase in rate of about 25, or, with a basal rate of 150, to an increase of about 17 percent. For stimulation by animals, the same figure shows in less than 1.25 sec increases of as much as 74 percent for cat *A* and 67 percent for cat *B*. Unless the comparison is invalid, it appears that the accelerators are not fast enough to account for the general type of acceleration secured from these cats.

Another of the curves of Rosenblueth and Simeone represents the temporal course of acceleration when combined accelerator and vagal stimulation changes to purely accelerator stimulation. The rise in 1 sec. is here from about 140 to 190, a rate of rise quite comparable to that produced with shot stimulation, although not as rapid as that involved in the 'spikes' of responses to other animals. In this case, however, the frequency of vagal stimulation was far from optimal. Presumably with increased frequency the subsequent rise would be considerably greater.

Finally, the two investigators give a number of curves showing the rate of decrease of the heart rate with stimulation of a depressor and with cessation of stimulation of an accelerator. Depressor retardation is distinctly more rapid than the rate of recovery from shot stimulation. Retardation from cessation of stimulation of the accelerator, however, is slower than the rate of recovery from shot: the former is about 10 percent of the base rate in 6 seconds, whereas Fig. 8 shows that the usual rate of recovery from shot stimulation is 4 sec and the usual extent of recovery, 25 beats per minute, or, using a basal rate of 145, about 17 percent. Here, again, the data suggest a marked participation of the depressor nerves in the emotional cardiac acceleration which we have studied. It seems legitimate to conclude that, in the cat, cardiac acceleration elicited by sudden stimuli—shots and large dogs—is largely effected by inhibition of depressor tone rather than by purely sympathetic excitation.

The second issue upon which our data bear is the validity of certain current arguments against the James-Lange theory, particularly of the one that "visceral changes are too slow to be a source of emotional feeling." [4] Our data show that in the cat an emotional stimulus elicits at least one visceral response which (*a*) can exhibit great uniformity for a specific stimulus (as in the case of a shot), which (*b*) can be of great magnitude (as when a dangerous dog is presented) and which (*c*) follows soon after the skeletal startle response, developing within .5 sec of stimulation.

Data from human beings would be more crucial in this connection than data from cats. We know of no specification of the temporal course of rise in heart rate elicited by shots in human beings, presumably because the response is less clear-cut than in cats. Lehmann, from a review of the older experiments (before 1914) concluded that "circulatory changes which are caused by simple sensory stimuli occur in general only 1 to 2 seconds after the beginning of stimulation, and consequently a little later than the feeling tone," [5] but his data appear very doubtful. Blatz reports that in a sudden fall, "The pulse rate was immediately accelerated from 88

per minute to 102 per minute." [6] Latency was not an issue, however, for Blatz, and it is not clear just what is meant by 'immediate.' It seems to mean 'before five seconds' rather than 'within one beat of stimulation.'

Fortunately there is some good evidence on the latency of change in form of the electrocardiogram in human beings. Landis and Slight found that in the twenty out of seventy-four cases where difference of form was manifested in the electrocardiogram after loud stimulation the difference appeared from .3 to .8 sec after stimulation, with an average of .48 sec.[7] It is likely, therefore, that in emotional situations the temporal course of the heart rate in human beings is about the same as in the cat. In this case, whatever may be the objection to the James-Lange theory on other grounds, the speed of visceral reactions need not militate against the theory.

Summary

This study deals with the changes in heart rate elicited in the cat by 'emotional' stimuli. Action potentials of the heart beat were amplified and recorded on ticker tape by means of an undulator. Stimuli consisted primarily of pistol shots and the presentation of dogs and other cats. A shot elicits a definite acceleration in heart rate, with a latency of .54 sec on the average. The action time (period of positive acceleration) varies, according to the magnitude of the response, from .42 to 2.3 sec. The recovery time also depends upon the magnitude of the response, but shows an average duration of about 4 sec. The extent of the rise in heart rate is about 30 beats per minute.

Sudden presentation of a strange and menacing dog produces an extremely marked change in heart rate. For the first two presentations to two cats the average rise was 81 percent of the base rate. In this rise there is usually a sharp peak in heart rate within .35 to 1.25 sec of stimulation.

The temporal features of the startle response to shots and the rapidity of occurrence of the sharp peak in the response to dogs may be compared to the temporal features of cardiac response to direct stimulation of autonomic nerves. The

comparison suggests that the sudden response to shots and dogs must be due to the activity of the parasympathetic system. These sudden responses show, furthermore, that criticism of the James-Lange theory on the basis of the slowness of visceral response finds no support in the behavior of the heart.

References

1. W. A. Hunt & C. Landis, Studies of the startle pattern, I, Introduction; II, Bodily pattern; III, Facial pattern, J. of Psychol., 1936, **2**, 201–219.
2. R. Hunt, Direct and reflex acceleration of the mammalian heart, with some observations on the relation of the inhibitory and accelerator nerves, *Amer. J. Physiol.*, 1899, **2**, 395–470.
3. A. Rosenblueth and F. A. Simeone, The interrelations of vagal and accelerator effects on the cardiac rate, *Amer. J. Physiol.*, 1934, **110**, 42–55.
4. W. B. Cannon, *Bodily Changes in Pain, Hunger, Fear and Rage*, 2d edit., 1929, 354. The quotation is the heading of a section. In the body of the section Cannon contends: (1) The viscera are composed of smooth muscles or glands, with the exception of the heart, which involves modified striate muscle; (2) smooth muscles or glands have latencies too great for the requirements of the James-Lange theory. Thus Cannon excepts the heart from his detailed argument. He does not, however, state that the heart reacts fast enough to fit the requirements of James' theory and the impression left with the reader is that *all* viscera are too slow. Freeman, for instance, writes: "Other evidence against the James-Lange theory includes . . . the long latency of visceral response," without excepting the heart. (G. L. Freeman, Introduction to Physiological Psychology, New York, 1934, 181.)
5. A. Lehmann, *Die Hauptgesetze des menschlichen Gefühlslebens*, Leipzig, 1914, 70.
6. W. E. Blatz, The cardiac, respiratory and electrical phenomena involved in the emotion of fear, *J. Exper. Psychol.*, 1925, **8**, 120.
7. C. Landis and D. Slight, Studies of emotional reaction, *J. Gen. Psychol.*, 1929, **2**, 413–420.

(Manuscript received January 9, 1937)

STUDIES ON THE PHYSIOLOGY OF SLEEP CHANGES IN IRRITABILITY TO AUDITORY STIMULI DURING SLEEP *

BY F. J. MULLIN, N. KLEITMAN, AND N. R. COOPERMAN

From the Department of Physiology of the University of Chicago

The basis for the construction of the classical curves of the 'depth of sleep' as found in the literature and textbooks of psychology and physiology is the work on the intensity of stimulus necessary to awaken a sleeping subject at various times during the night. Such studies were made by Kohlschutter (**5**), 1862, using the sound produced by a pendulum falling from a height onto a slate plate; by Monninghof and Piesbergen (**8**), 1883, who employed a lead ball falling on an iron plate; by Michelson (**7**), 1891, with the use of various brass balls falling from a given height to a table; and by Czerny (**1**), 1891, using induction shocks. These so-called curves of the intensity of sleep all show a peak during the first or second hour of sleep followed by a drop to a low level. In some of these and also in other early curves obtained by similar methods, particularly that of Lambranzi (**6**), 1900, there occurs a later peak of about half the height of the first, coming near the end of the sleep period. In 1923 Hass (**2**), stimulating with a sound, found that the intensity necessary to awaken was very great thirty minutes after the onset of sleep and was considerably less during the remainder of the night.

In 1933 Kleitman, Cooperman and Mullin (**4**) reported motility changes during sleep which, when plotted on an hourly basis, gave a curve roughly inverse to those of the depth of sleep. On the other hand, the curve of the distribution of periods of immobility of half an hour or longer in

* This work was aided by a grant from the Rockefeller Foundation. Thanks are due to Messers. N. Brewer, A. T. Doktorsky, M. Hanley, N. Morris, Y. T. Oester, L. Palitz, I. Rossman and S. Titelbaum for their valuable assistance in obtaining the data reported.

duration during the night bears a close resemblance to these classical curves of the depth of sleep. The similarity of these two types of curves led to the development of a working hypothesis that the apparent depth of sleep as determined by the intensity of the stimulus necessary to awaken the sleeper might be a function of the length of time that elapsed after he had changed his position.

Methods

This study consists of two parts: the determination of the changes in irritability to auditory stimuli which occur in a dog when it goes to sleep, and the determination of the intensity of auditory stimuli necessary to awaken a human sleeper at various times during the night in relationship to his motility. In both cases the auditory stimulus was delivered by a magnetic loudspeaker connected to a 110-volt A.C. outlet through a rheostat. The rheostat dial was graduated from 0 to 100, and the voltage represented by the figures on the dial was calibrated by means of a voltmeter. It was thus possible to grade the intensity of current flowing to the loud-speaker, and to use this as a measure of the loudness of the sound emitted by the latter.

In the dog experiments the animal was placed in a stand which held its head in place and which could easily be adjusted to the height of the dog. A wide band of ticking was passed underneath the dog's body as a sort of sling, which was loose enough so that if the dog stood on its own feet the sling would pass freely under its belly, but the dog had only to relax a bit and the sling would then support it. The legs of the dog were entirely free. The dogs had been taught to remain in the stand quietly. The operator would stay in an adjoining room and observe the animal continuously through a peep-hole in a covered glass window so arranged that the dog could not see he observer. The main criterion as to whether the dog was asleep or not was the observation of its postural attitude. Under the conditions of these observations the dogs would get drowsy, as evidenced by a relaxed body entirely supported by the sling with the head slumping down and the eyes closed, and would go to sleep for various intervals of time.

After some experience it became fairly easy to judge as to when the dog was dozing or asleep and when it was awake. When the observer entered the room of the sleeping dog for any purpose, the animal could often be heard snoring and would wake with a start if disturbed. Records were kept of the behavior of the dog while in the stand and the time when it went to sleep and when it awoke and what the apparent cause of its awakening was. Some dogs would not sleep longer than a minute or two at a time even after three months of daily trials, and such dogs were not used. The dogs that readily went to sleep in the stand were stimulated by the sound emitted from the loud-speaker at various intervals of time after going to sleep, and the intensity of stimulus necessary to awaken the dog was plotted against the length of time the dog had been asleep.

In the work on human beings the subject slept in a bed so arranged that a continuous kymographic record of his motility, as represented by movements of the bedspring, could be obtained each night (See Kleitman, Cooperman and Mullin (4), 1933). The observer was in a non-adjacent room where a light would shine every time the subject moved during the night. When making a determination the experi-

menter actuated the loud-speaker which was located in the subject's room a few feet from the head of his bed. This auditory stimulation was recorded by a signal magnet on the kymograph tracing of the sleeper's motility. If the subject was awakened by the sound of the loud-speaker, he pressed a signal key which sounded a buzzer in the experimenter's room and at the same time made a signal magnet record in the kymograph set-up. It was thus possible for the experimenter to stimulate the sleeper at any time during the night, either with or without reference to the latter's movements, and to grade the intensity of current going into the loud-speaker. When stimulating the sleeper the following routine procedure was employed. The experimenter turned the rheostat dial to 10 and produced that sound for ten seconds. If the subject did not respond, the rheostat was rapidly turned back to 0 for ten seconds and then turned up to 20 and held for another ten seconds, and so on until the sleeper was awakened enough to press the signal key, or until the maximum obtainable loudness was reached.

Results

On the basis of over one hundred and fifty trials on five dogs the irritability to auditory stimuli was plotted against time after the onset of sleep (Fig. 1). For each of the dogs

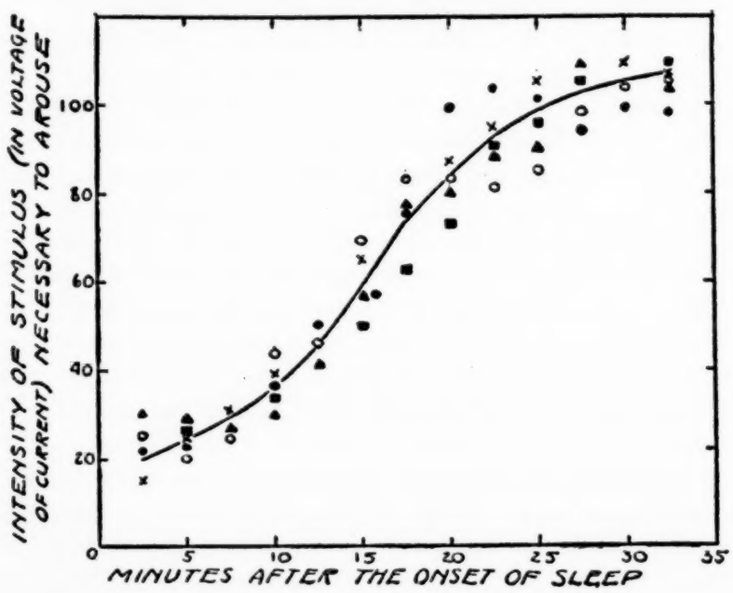

Fig. 1. Irritability changes in dogs following the onset of sleep. The average of the determinations for each of five dogs is indicated by an appropriate mark on the graph at the various time intervals indicated. The curve represents the average of all the dogs together, and records the rise in the threshold of response to sound. This curve is constructed from a total of one hundred and sixty-three determinations.

the result was a curve somewhat S-shaped in character and reaching its highest level in from thirty to forty minutes. The change from the concavity to the convexity in the curve occurred about fifteen minutes after the onset of sleep. Sufficient experimental observations were not obtained to

carry the curve beyond its highest level, but some observations made indicate a gradual return toward the waking threshold during the half-hour following the point of greatest depth of sleep. A few tests were also made on the irritability changes in human beings on first going to sleep. The experimental procedure involved direct observation of the subject by the observer, who stimulated the sleeper at various intervals, after the onset of sleep, in the usual manner with a loud-speaker. The results show a curve of decreasing irritability for the human subjects similar to that seen in the case of the dogs.

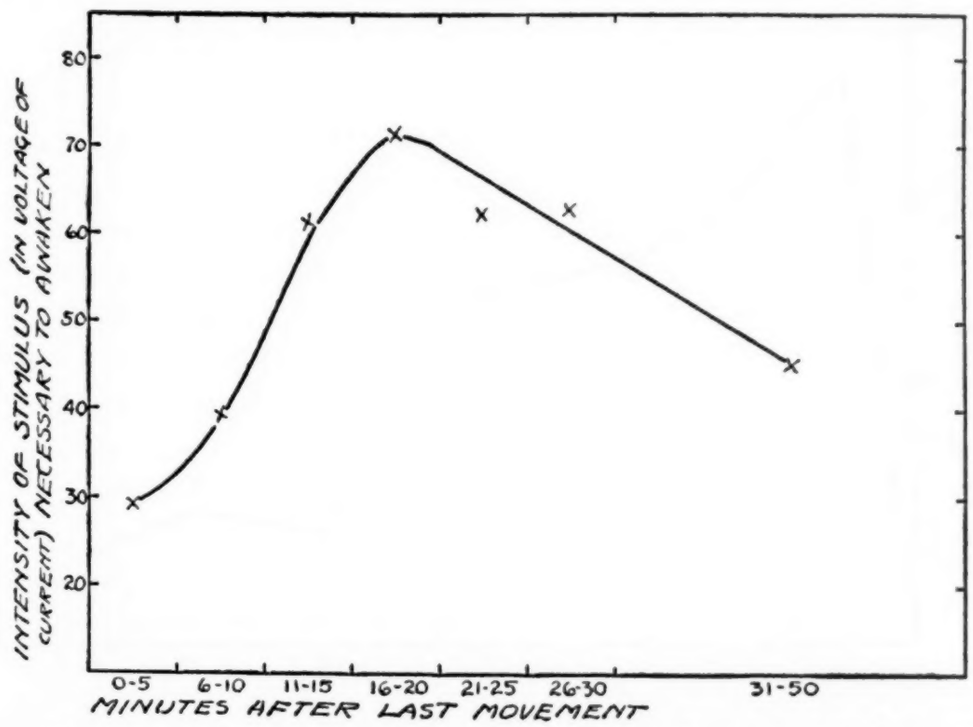

FIG. 2. Irritability changes in human subjects during nocturnal sleep. This curve shows the relationship of the irritability, as measured by auditory stimulation, to the length of time since the last movement of the sleeper. The plotted points represent the average threshold of response of six male subjects as determined by a total of three hundred and twenty-five observations.

The relationship of the average intensity of auditory stimulus necessary to awaken the human subject during the night's sleep to the length of the period of immobility immediately preceding the stimulus was worked out for six male subjects by some three hundred tests. This can be represented by a curve (Fig. 2) which is slightly S-shaped in its

ascent. The change from the concavity to the convexity in this curve occurs in approximately ten minutes, and the height of the curve is reached between sixteen and twenty minutes of immobility. Thereafter the curve of the threshold of response shows a downward direction, indicating shallower sleep. With a few exceptions, the longer the interval of quiescence between pairs of successive movements the higher the intensity of stimulus necessary to awaken, and the more prolonged is the decrease in irritability, no matter what time of night this occurred.

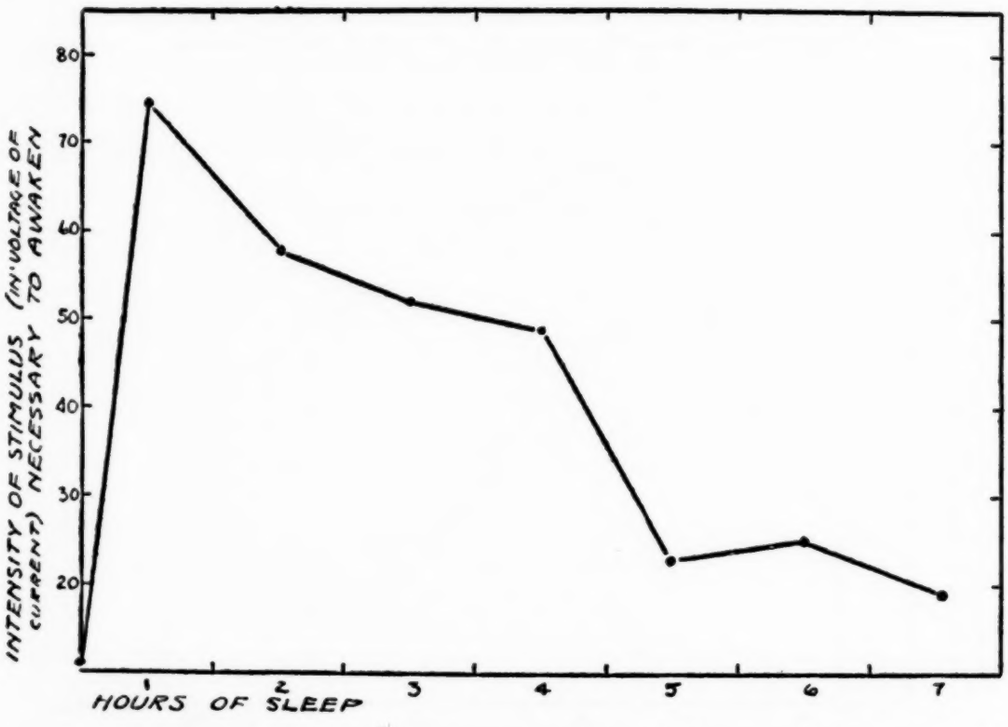

FIG. 3. Irritability changes in human subjects during a night's sleep. This curve is constructed from the same experimental data used for Fig. 2, and each point represents the average change in irritability for successive hours.

If the data that were used for the construction of the curve for the intensity of sound necessary to awaken the sleeper in relation to the time following a movement are now plotted with respect to the hour of the night during which the subject was stimulated, averages are obtained from which a composite all-night curve of the depth of sleep can be constructed (Fig. 3). This curve is none other than the classical depth of sleep curve based on the averages of the

results obtained by random sampling throughout the night on different occasions.

The amount of movement during the night and also the average intensity of stimulus necessary to awaken varied considerably in the different individuals studied. Each of these factors seemed to be characteristic for a given sleeper, but the differences were not always in the same direction. Thus, the most quiet sleeper was also the one most easily aroused, on the average. The same types of motility curve and irritability curves were obtained on all subjects, however.

Discussion

The consistency of the results obtained and the fact that experimental observations repeated at various times checked each other quite well leads one to believe that a real and variable factor in association with sleep is being adequately measured by this method. Some earlier unpublished work on dogs done before the present investigation was started gave indications that a stronger induction shock had to be applied to the leg of the animal some minutes after it had fallen asleep in the stand than more immediately following the onset of sleep. The increase in the depth of sleep as measured by the induction shock method resembles that obtained by auditory stimulation (Fig. 1), indicating the general character of the phenomenon. What the exact significance of the S-shape in the ascending portion of the curve of depth of sleep is remains to be elucidated.

The ascending portion of the curves for dog and for man are similar. The fact that the maximum intensity of sleep in the dog is reached later may be due to a species difference, but may also be due to selective sampling, in that those for the dogs were determined at the beginning of a sleep period whereas the figures for the human subjects were obtained at various hours of the night. The latter are therefore composite figures representing the average of curves of longer and shorter duration.

Following any major type of movement there is a gradual increase and then decrease in the threshold of response to

auditory stimuli. Thus, for human subjects a series of changing irritability curves during successive periods of immobility throughout the sleeping period is more characteristic than the previously elaborated single curve for intensity of sleep. The early studies which yielded the classical curves of the depth of sleep are fully discussed and criticized in considerable detail by Johnson and Swan (**3**), 1930, who point out the inadequacy of the experimental data and the uncritical nature of some of the conclusions drawn therefrom. Several of these early workers found that it was often harder to awaken a sleeper after he had recently stirred and relapsed back into quiet sleep than at ordinary times. But such results were often discarded for one reason or another. It is of interest that the addition of a record of the sleeper's motility and the correlation between irritability changes and the duration of the periods of quiet sleep show that the irritability changes during sleep are associated rather directly with motility and quiescence, both factors showing similar variations during the night. Each person seems to have his own pattern of motility and level of irritability variations which he tends to follow more or less regularly.

In a normal night's sleep the greatest motility as well as the most frequent spontaneous awakenings of the sleeper fall in the last few hours of sleep. Toward morning the sleep becomes shallower and shallower, as reflected in the greater ease with which the person can be wakened at this time. Occasionally, however, a period of quiescence lasting forty-five minutes or longer will occur toward the end of the night's sleep, and in such cases the depth of sleep curve is similar to those found with the same duration of immobility in the earlier part of the night. This undoubtedly accounts for the secondary peaks found in some of the curves of the depth of sleep reported by earlier workers. Some of the earlier observations on the effects of alcohol and caffein on the intensity of sleep, as measured by Monninghof and Piesbergen (**8**), 1883, and also by others since, can probably also be accounted for on the basis of the motility changes caused by these drugs. Mullin, Kleitman and Cooperman (**9**), 1933,

reported that the relative frequency of motility and spontaneous awakenings during the night after taking these drugs shows that alcohol decreases, and that four and six grain doses of caffein increase the activity of the sleeper.

A person may wake several times during the night and press a signal key, and yet not recall having done so the next morning. A movement is nearly always associated with such spontaneous awakenings of the sleeper, which occur much more frequently in the last part of the night, but usually the person does not remember on the following morning exactly how many times he signalled. Toward the end of the night's sleep period increasing internal and external stimuli probably combine to make the periods of quiet sleep progressively shorter in duration until finally the sleeper wakes for the day, his irritability level having returned to the waking state.

Summary

1. A method for studying irritability to sound at the onset of and during sleep is described.

2. When a dog falls asleep, the threshold for auditory stimuli rises in a characteristic way and shows an S-shaped curve which reaches its maximum about half an hour after the onset of sleep.

3. A similar curve of irritability changes in the human sleeper is obtained when one plots the intensity of auditory stimulus necessary to awaken against the time after the last movement.

4. Changes in irritability during sleep in the human subject are directly related to the time that has elapsed since the last movement, rather than to the time of night.

5. Each person seems to have a characteristic motility pattern and irritability changes during sleep.

6. The classical curve of 'depth of sleep' does not adequately express the facts relative to changes in irritability during sleep; instead of one continuous curve of intensity of sleep there are many contiguous curves between successive movements, these curves getting briefer and shallower as the night progresses.

(Manuscript received November 27, 1936)

References

1. Czerny, A., Beobachtungen uber den Schlaf im Kindesalter unter physiologischen Verhaltnissen, *Jahrb. f. Kinderheilkunde*, 1891, **33**, 1–28.
2. Hass, A., Uber Schlaftiefenmessungen, *Psychol. Arb.*, 1923, **8**, 228–264.
3. Johnson, H. M. and Swan, T. H., Sleep, *Psych. Bull.*, 1930, **27**, 1–39.
4. Kleitman, N., Cooperman, N. R., and Mullin, F. J., Motility and body temperature during sleep, *Amer. J. Physiol.*, 1933, **105**, 574–584.
5. Kohlschutter, E., Messungen der Festigkeit des Schlafes, *Z. rationelle Med.*, 1862, **17**, 209–253.
6. Lambranzi, R., Sulla profondita del Sonna, *Rev. spreimentale Freniatria*, 1900, **26**, 828–830.
7. Michelson, F., *Untersuchungen uber die Tiefe des Schlafes* (Dissertation). Dorpat, Schnakenburg, 1891.
8. Monninghof, O. and Piesbergen, F., Messungen uber die Tiefe des Schlafes, *Z. Biol.*, 1883, **19**, 114–128.
9. Mullin, F. J., Kleitman, N., and Cooperman, N. R., The effect of alcohol and caffein on motility and body temperature during sleep, *Amer. J. Physiol.*, 1933, **106**, 478–487.

BLOCKING IN MENTAL AND MOTOR TASKS DURING A 65-HOUR VIGIL

BY NEIL WARREN AND BRANT CLARK

University of Southern California

Introduction

Studies of prolonged vigils and periods of reduction in sleep have been conducted by numerous investigators to determine both psychological and physiological reactions. The conclusions from these investigations have been largely negative as far as the test results are concerned, and where the results have been positive, they were not in proportion to the loss of sleep. There are several reasons for this. The normal sleep ration undoubtedly leaves a very large safety factor over and above the actual needs of the individual, and the test methods are not adequate to show reduction in efficiency for the reason that the subjects are able to compensate during the relatively short periods of the tests. However, anyone who has gone without sleep for 48 hours or more can testify that his subjective reactions do not correspond to the test results, and the introspections of experimental subjects are also definitely in contrast with these results. Without exception, the subjects reported that they felt tired, desired sleep, had to put forth additional effort, and so forth. No attempt will be made to review previous findings except as they bear directly on the results of this study. The studies of Moss (**5**), Weiskotton (**6**), and Cooperman (**4**), give extended discussions of earlier experiments. It is worth noting, however, that most of the studies, with the exception of Weiskotten's (**6**), have been concerned with tests before and after a prolonged vigil and have been less concerned with changes during the vigil.

Bills (**1**, **2**, **3**) has made several studies of blocking during fatigue. He uses the term 'block' to refer to those periods experienced by mental workers when they seem unable to

make a response and cannot continue, even by effort, until a short time has elapsed. He has arbitrarily defined a block as 'a pause in the responses equivalent to the time of 2 or more average responses.' He believes that these blocks tend to eliminate the decrement in fatigue, correspond to the rhythm of attention, and are associated with the refractory phase.

Bills measured blocks during alternate addition and subtraction, reversible perspective, color naming, naming opposites, and substitution. The subject made the responses orally, and one experimenter pressed a key which recorded the response on kymograph paper. Another recorded errors in the same way. The subjects were instructed to work as fast as possible, consistent with accuracy. Bill's results may be summarized briefly as follows: In mental work involving considerable homogeniety and continuity, blocks occur with almost rhythmic frequency. Fatigue increases the size of the blocks, and practice reduces them. Individuals who respond rapidly have fewer blocks. Errors occur in conjunction with blocks suggesting that the cause of the errors lies in recurrent low neural functioning which the blocks reveal. The work curves for blocking showed much larger decrement than those for response rate. He also found that stutterers block about twice as often as normals. In general, Bills believes that his results confirm the theory that blocks increase the amount of rest gained and hence enable the subject to maintain his speed and compensate for fatigue.

Whiting and English (7) have advanced the hypothesis that fatigue is a negative emotional appetite and should be differentiated from the physiological phenomenon of exhaustion. In this sense, fatigue is a 'negative motive to action' and does not directly cause a work decrement but raises the threshold at which certain work motives are effective. On this basis, if the positive motives are at all adequate, the fatigue will have little effect on work efficiency. This explanation of fatigue, as contrasted with exhaustion, on the basis of motives, fits in quite well with the results found during prolonged sleeplessness and will be discussed further in relation to the results of this study.

Experimental Procedure

The experimental vigil lasted approximately 65 hours, from 5:30 A.M. Friday morning until 10:30 P.M. Sunday night. Four students in psychology acted as subjects. All were paid for their time as an added motivation. A complete physical examination was given each subject before the test and on the morning of the last day. Three of the subjects completed the experiment, and the fourth completed approximately 48 hours. Preliminary tests were made on three of the S's before the vigil began and post-tests were given after the experiment was completed. The preliminary test was not given the fourth subject because he volunteered to act as a subject the day that the test began. The tests were begun at 6:00 A.M., and the complete battery was given in the morning. On Friday (the first day), the subjects were permitted to attend classes as usual, but from mid-night until the completion of the experiment, they were not permitted to be alone at any time.

Various methods were used to keep them awake, *e.g.* games of various sorts and walks were particularly effective. The tests were made frequently enough to keep them occupied a large part of the time. Experimenters and assistants worked in shifts. A majority of the tests were made by the authors, but three graduate students made some of them. About 12 assistants worked in shifts keeping the subjects awake. Near the end of the testing period, it was impossible to eliminate naps entirely. One subject had a fifteen minute nap Sunday morning, but with this exception the naps were limited to a few seconds. Another subject declared himself unable to continue after being awake for 48 hours. He was given a final series of tests and permitted to leave. All of the subjects coöperated well throughout the testing period. The tests were given at ten hour intervals on a regular pre-arranged schedule.

The results to be reported in this paper cover only a part of the tests which were used. Aside from three tests for blocking, the following were given: Phonograph records of speech; binocular eye-movement photography of reading, convergence, and simple interfixations; steadiness, aiming, writing; a series of optometric tests, peripheral vision, color vision, depth perception; and free association. The results of these tests will be reported elsewhere.

Three tests of blocking were made, two of which were similar to those of Bills. The tests were: alternate addition and subtraction, color naming, and tapping with the index finger of the right hand. The experimental procedure was similar to that used by Bills (1). The subjects were instructed to work as rapidly as possible consistent with accuracy, and as the subject responded orally, one experimenter recorded the response by pressing a key which marked on a kymograph drum. Errors were recorded by a second experimenter. A time line marking seconds was also recorded.

The addition and subtraction involved the alternate addition and subtraction of 3 from 384 intergers, which ranged from 4 to 9. The subjects were allowed 3.5 minutes to complete the series but were unable to do so every time. The following six colors arranged in random order were used for the color naming experiment: Red, orange, yellow, blue, green, purple. There was a total of 140 colors, and the subjects were permitted one minute to complete the series.

In tapping, the subject tapped a key on the downward movement and also touched a bar 1.3 inches above the key before beginning another tap. Two trials of one minute each were taken.

All of these tests of blocking were made at one sitting. Addition and subtraction was measured first, color naming second, and the tapping last. The subjects were permitted to rest between the two tapping trials.

Since the present technique varied slightly from that of Bills, a control series of tests was made on four new subjects who were not required to go without sleep. They were measured on the same tests at intervals of 12 hours for a period of 60 hours. The conditions were comparable to the experimental procedure except that all tests were made during the daytime and a normal amount of sleep was permitted.

Results

For the purpose of this study a block was arbitrarily defined as a failure to respond equal in time to two or more modal responses. Variations from the mode were more satisfactory measures than from the average since the latter was more influenced by changes in the number of blocks.

The data were stated in terms of the percentage of total responses which were blocks as determined by this criterion.

Addition and Subtraction.—The composite percentages of the four experimental subjects (Fig. 1) shows only a slight increase in the proportion of blocks during the first 24 hours of sleeplessness. Two of the four individual records show slight decreases in blocking at the 20-hour test; one had a slight increase, and the fourth subject had more than doubled in number of blocks from his first to his third test (see Table 1).

TABLE 1

Composite Records of Experimental Subjects (Addition and Subtraction)

	Hours of Sleeplessness						
	0	10	20	30	40	50	60
Number of responses	357	354	381	384	325	345	338
Modal response (sec.)	.63	.66	.62	.54	.53	.59	.55
Percentage of blocks	9.2	9.9	12.5	14.0	23.5	21.4	18.5
Average block (sec.)	1.9	1.5	1.5	1.4	1.4	1.6	1.4
Percentage of errors	3.6	2.1	1.1	0.6	1.9	2.1	1.6

The average percentage of blocks in addition and subtraction for the initial trial on the morning of the first day was 9.2, ranging from 6.4 percent to 12.7 percent. The peak in the number of blocks was reached at the 40 hour test for the group as a whole, the average having increased to 23.5 percent. The individual records show a variation from the average with an increase in blocking after sleeplessness indicated in each case. Subject L, who was unable to

complete the vigil, had an increased number of blocks from 6.4 percent on his first trial to 27.3 percent on his final trial just before he was taken home at the end of 48 hours. Subject T had increased from 12.7 percent blocks on his first test (5.5 percent on the second) to 32.4 percent after 40 hours of sleeplessness. In the case of Subject S, 24.7 percent on his total responses were blocks after 50 hours, as compared with 8.5 percent on his first trial. His peak, however, was reached 10 hours later on the next to the last trial with 31 percent. Subject F was relatively constant in number of blocks during the earlier tests, starting with 9.3 percent, reaching a low point of 2.2 percent at the 40 hour test and a high point on the final test with 27.1 percent. The other two subjects showed a decrease in number of blocks on the final trial, apparently indicating an increase in motivation upon nearing the completion of the vigil. In each case the final test percentage was above that of the initial trial and the average of 18.5 percent was double the first average.

The four control subjects, on the other hand, started with a higher proportion of blocks on the first tests, remained relatively constant throughout until the last trial in which each showed a marked decrease in number of blocks. The average percentage of blocks in addition and subtraction was 6.7 for the final test, compared with 17.2 percent for the first test (see Table 3). At no time was the proportion of blocks appreciably above the first trial.

There was no significant change in length of blocks during the experimental period (Table 1). This was also true for color naming (Table 2).

TABLE 2

COMPOSITE RECORDS OF EXPERIMENTAL SUBJECTS (COLOR NAMING)

	Hours of Sleeplessness						
	0	10	20	30	40	50	60
Number of responses	136	131	127	116	103	130	126
Modal response (sec)	.49	.45	.43	.43	.40	.35	.40
Percentage of blocks	12.8	14.2	14.5	19.5	21.8	23.1	16.2
Average block (sec)	1.2	1.1	1.2	1.2	1.2	1.2	1.0
Percentage of errors	1.1	0.8	0.8	1.1	1.3	1.2	0.6

TABLE 3

COMPOSITE RECORDS OF CONTROL SUBJECTS (ADDITION AND SUBTRACTION)

	Hrs. of Interval					
	0	12	24	36	48	60
Number of responses	326	347	322	340	325	349
Modal responses (sec)	.59	.49	.56	.46	.47	.52
Percentage of blocks	17.2	18.3	13.2	17.1	17.7	6.7
Average blocks (sec)	1.3	1.2	1.2	1.1	1.2	1.2
Percentage of errors	3.3	2.1	1.0	1.3	1.5	.8

Color Naming.—The results for color naming test were consistent with those for addition and subtraction. The experimental subjects started with 12.8 percent blocks on the first test, reached a peak of 23.1 percent on the 50 hour test and improved on the last test. The result for the last trial was 16.2 percent blocks. For each subject the greatest number of blocks occurred on his last or next to last trial.

The averages for the control subjects remained fairly constant throughout, and the percentages of three of the four were smaller on the final test than on any of the earlier ones. The records for the group are given in Table 4.

TABLE 4

COMPOSITE RECORDS OF CONTROL SUBJECTS (COLOR NAMING)

	Hrs. of Interval					
	0	12	24	36	48	60
Number of responses	106	122	116	118	126	126
Modal response (sec)	.38	.37	.39	.35	.31	.34
Percentage of blocks	13.6	17.1	13.9	14.	14.1	11.4
Average blocks (sec)	.9	.9	.9	.8	.8	.8
Percentage of errors	.9	1.3	1.3	.6	.2	.4

Errors.—There was no apparent relationship between loss of sleep and number of errors in addition and subtraction nor in color naming. Subject T, who had no preliminary training trial, had 26 errors on his first trial in addition and subtraction but the others were relatively constant throughout varying from 0 percent to 5.1 percent of the total responses in addition and subtraction, and from 0 percent to 3.2 percent in color naming.

A relationship of errors to blocks is apparent since 72.3 percent of the errors in addition and subtraction and 76.2 percent in color naming are associated with blocks. In the greater proportion of these instances the block followed the error, indicating that the error may have acted as a distraction which caused the block. This was true for both experimental and control subjects.

Tapping.—It has been pointed out that in tapping tests there are frequent 'breaks' or blocks during which no response is made. These pauses resemble the blocks that occur in mental work of the kind already described. For that reason it might be indicated that both types of blocks result from the same causes. In the present study, however, tapping blocks did not occur with the frequency of blocks in the mental tasks. Neither was there an increase in proportion of blocks after extended sleeplessness.

Blocks in tapping were more likely to occur during rapid tapping than after the subjects had begun to slow down as a result of finger fatigue. This seems to indicate that as the subject took a steady pace, attending to each tap, there was less tendency to break in the rhythm.

Each of the experimental subjects had a decrease in total number of taps, as well, as of taps per second, after 30 hours of sleeplessness. There was an increase in each case on the final trial. The average total number of taps was 328 for the first trial, dropped to 186 at the 30 hour test, and was back to 306 on the final trial. The average number of taps per second for these three intervals was 3.9, 3.2, and 4.2. Evidently there is a decrease in motivation or, perhaps, an increase in 'negative motivation' as a result of sleeplessness which is overcome toward the end of the vigil.

Conclusions

Blocking in mental work was greatly increased after a period of prolonged sleeplessness, the most noticeable increase being observed after 40 hours for all experimental subjects. This was true in spite of the fact that the modal response did not change appreciably, and the average response changed

but slightly. Another indication of blocking was observed in all subjects and in experimenters and assistants who had very little sleep during the vigil. There was a universal difficulty in recalling names of persons or objects mentioned in conversation.

The final test period indicated a decrease in blocking for two of the three subjects who completed the vigil. Where this final 'spurt' occurs it is evident that measurement at the beginning and at the end of a period of sleeplessness does not give a true picture of the changes produced by the vigil.

As Bills has suggested, the occurrence of blocks seems to be related to variations in the process of attending. This is supported by the association of errors with blocks, particularly when the block follows the error.

It is obvious that the sleeplessness was not continued to the point of exhaustion. The decrement which occurred seems best explained, as Whiting and English have suggested, on the basis of failure to achieve motivation sufficient to overcome the subjectively increased thresholds of attention and effort. In a number of test situations the desire to stop trying became more noticeable as the sleeplessness was continued.

In contrast with the mental work there was no increase in the number of blocks during tapping, and the blocks were fewer throughout. This is possibly to be explained as due to the fact that the subject was not motivated to work beyond a congenial pace.

If the results of the increase of 'negative motivation' during prolonged sleeplessness are to be measured, the experimental procedures must take into account the following three points:

1. Tests should be given at regular intervals during the sleepless period to show changes in motivation.

2. Tests should be of sufficient length to diminish the effect of compensation.

3. Where possible tests should be devised which make compensation difficult or impossible.

(Manuscript received December 7, 1936)

References

1. Bills, A. G., Blocking: A new principle of mental fatigue, *Amer. J. Psychol.*, 1931, **43**, 230–245.
2. Bills, A. G., Some causal factors in mental blocking, J. Exper. Psychol., 1935, **18**, 172–185.
3. Bills, A. G., The relation of stuttering to mental fatigue, J. Exper. Psychol., 1934, **17**, 574–584.
4. Cooperman, N. R., Mullin, F. J., and Kleitman, N., Studies on the psychology of sleep, XI, *Amer. J. Physiol.*, 1934, **107**, 589–593.
5. Moss, F. A., *et al*, The measurement of fatigue by physiological methods, J. Exper. Psychol., 1931, **14**, 423–438.
6. Weiskotten, T. F. and Ferguson, J. E., A further study of the effects of loss of sleep, J. Exper. Psychol., 1930, **13**, 247–266.
7. Whiting, H. F. and English, H. B., Fatigue tests and incentives, J. Exper. Psychol., 1925, **8**, 33–49.

A COMPARISON OF COLLECTIVE AND INDIVIDUAL JUDGMENTS OF FACT

BY HERBERT GURNEE

Western Reserve University

It is known that collective adjustment to a maze situation is likely to be significantly more efficient than average individual adjustment.[1] The question arises whether similar superiority of collective action would obtain in other kinds of situations. Studies of relevant problems make this seem probable. Pooling of individual judgments is known to yield, under some conditions at least, values superior to the average.[2] There is also evidence that, after group discussion, individuals may manifest greater accuracy of report of perceived material [3, 4] and a more intelligent viewpoint on ethical questions [4] than previously. Moreover, it has been shown that participation in a committee or jury tends to produce more correct solutions to mental problems [3, 5, 6] and a more accurate account of a reported event [7] than individual action, although it appears that the differences are sometimes insignificant.[8]

These investigations, for the most part, were concerned primarily with the effects of discussion, either upon the

[1] GURNEE, H., Maze learning in the collective situation, *J. Psychol.*, 1936, 3, 437–443.

[2] GORDON, K., Group judgments in the field of lifted weights, J. EXPER. PSYCHOL., 1924, 7, 398–400. Also H. C. KNIGHT in G. and L. Murphy, *Experimental Social Psychology*, 1931, pp. 504–506.

[3] JENNESS, A., The rôle of discussion in changing opinion regarding a matter of fact, *J. Abn. and Soc. Psychol.*, 1932, 27, 279–296.

[4] BECHTEREW, W. and DE LANGE, M., Die Ergebnisse des Experiments auf dem Gebiete der kollectiven Reflexologie, *Zsch. f. angew. Psychol.*, 1924, 24, 305–344.

[5] SHAW, M. E., A comparison of individuals and small groups in the rational solution of complex problems, *Amer. J. Psychol.*, 1932, 44, 491–504.

[6] WATSON, G. B., Do groups think more effectively than individuals? *J. Abn. and Soc. Psychol.*, 1928, 23, 328–336.

[7] DASHIELL, J. F., in *Handbook of Social Psychology*, 1935, pp. 1135 ff.

[8] BURTT, H. E., Sex differences in the effect of discussion, J. EXPER. PSYCHOL., 1920, 3, 390–395.

individual participant or upon the achievement of a small group. Since discussion complicates the conditions of a social event, it would seem desirable, as a further step, to eliminate this variable and observe the effectiveness of collective action in the absence of all verbal interstimulation. Such procedure simplifies the situation beyond what normally obtains, but simplification is a common necessity of experimental analysis.

Specifically, the problem of the present study was to compare individual and collective judgment on the truth or falsity of various statements, under conditions which precluded discussion and other means of communication. The individual judgments were obtained by the procedure usually employed in true-false tests. The collective judgment was obtained as in ordinary parliamentary practice, namely, by a majority vote of the group on each statement submitted. In order to secure reliable results, it seemed advisable to use groups larger than the ordinary committee or jury.

Procedure

For the purpose of sampling collective judgment upon different kinds of facts, three experimental situations were used. These situations were presented in all to five different groups, varying in size from 18 to 66 subjects.

Situation 1.—The aim here was to submit statements about facts that all subjects had the same opportunity and were under the same compulsion to learn. Accordingly, an undergraduate class, numbering 53 subjects, was given one of their regular monthly tests. This test consisted of 50 true-false statements on material assigned in the textbook. The usual individual responses were obtained and the papers collected; then, after necessary instructions, the statements were read one at a time by a member of the class and a vote was taken. Voting was by acclamation, with a show of hands in cases of an apparent tie. When, occasionally, a quorum did not appear to be acting, a second vote was required. The experimenter put the question each time and recorded the response. Subsequently, the individual tests were graded and comparison made with the group responses. Later a similar procedure was carried out with a different list of questions on a different group.

Situation 2.—Here the purpose was to submit statements about matters concerning which there were presumably varied experiences and interests. Accordingly, a chemical information test of 60 true-false statements was prepared. The facts contained in the test concerned various chemical matters of an essentially non-technical sort, ranging from simple to very difficult. This test was given to a class of 66 persons. Individual responses were first obtained, then collective reactions, in the same manner as described in Situation 1.

Situation 3.—Here statements were submitted about a complex social event which all had just had the opportunity to observe. The event consisted of a motion-picture

sequence of about three minutes duration. The film had been prepared by Metfessel [9] and showed a group of officers pursuing and capturing two criminals who were trying to escape in an automobile. The action was rapid and complicated. After seeing the film, each member of the class, which numbered 18 persons, was asked to check a list of 45 statements about the contents. The papers were collected and a vote was taken by show of hands on whether each item was 'true,' 'false' or 'did not appear.' The situation was repeated with another group of 18 persons.

Results and Discussion

The groups in every situation not only excelled the average individual performance in number of correct judgments, but equaled or approximated the performance of the best member. The results are thus comparable to those reported previously on collective learning of a maze problem. The table presents the comparative data for the three situations.

	Number of Right Responses				
	Situation 1		Situation 2	Situation 3	
	A (n53)	*B* (n57)	*C* (n66)	*D* (n18)	*E* (n18)
Group vote	50	48	57	28	30
Highest indiv.	48	49	57	27	32
Second indiv.	46	47	56	27	29
Indiv. Av.	35.0	37.4	47.4	23.3	24.5
Indiv. S.D.	6.7	5.1	5.0	2.7	3.1

A, *B*, etc., represent the different aggregations or groups of subjects. Fifty is a perfect score in *A* and *B*, 60 in *C*, and 45 in *D* and *E*. By chance alone the expected scores would be 25, 30 and 15 in the three respective situations. In only two of the five instances does a member exceed the group performance, and then by an insignificant margin. The second highest member is below the group in each instance. These comparisons are especially interesting for *C*, where the two leaders each had five semesters of chemistry, whereas two-thirds of the group had two semesters or less. The size of the standard deviations indicates that the group superiority over the average individual response is reliable in every case.

[9] Metfessel, M. and Musgrave, Jr., H., *Instructors Guide for Demonstrations of Psychological Experiments*, 1933, pp. 51 and 77.

Since the group result is necessarily a function of individual judgments, collective achievement is, as the data show, a relative and not an absolute thing. A situation which is difficult for most of the participants will be difficult for the group. Thus groups *A*, *B* and *C* are 95 to 100 percent correct in their responses, but groups *D* and *E*, confronted by a different sort of situation, are only 62 and 67 percent correct. Accordingly, while one can predict, in circumstances such as these, that the group will definitely excell the average performance of its members, he can not predict the accuracy of the group without knowing what the individuals themselves would do.

Perhaps the best way to express the group performance is to state it in terms of sigma distance from the average performance of the individual members. If we do this, we find that group *A* is $+ 2.2\sigma$ from its average, *B* is $+ 2.1\sigma$ from its average, *C* is $+ 1.9\sigma$ from its average, *D* and *E* are both $+ 1.8\sigma$ from their averages. These values are all strikingly similar, and suggest a rather constant psychometric function. Venturing a tentative statement of such a function, we may say something like this: By simple collective action the frequency of correct judgments of a group of persons will be increased approximately plus two sigmas over the average of their individual responses. Only further investigation can determine whether such a function generally holds, and, if it holds, to what extent it may be affected by such variables as discussion, suggestion, mental set, size of group, and so on. Needless to say, a superiority represented by plus two sigmas would mean that only about one member in forty could be expected to equal or exceed the group performance.

How is this superiority of collective judgment to be explained? Analysis of the individual responses reveals that the predominant factor was a distribution of individual errors. Taking group *A*, we find that a majority of the individual responses, obtained just before the collective action, were right on 46 of the 50 questions. Similarly, in group *B*, a majority were right on 43 out of 50 questions, and, in *C*, a majority were right on 52 out of 60 questions. In *D* and *E*,

a plurality of individuals were right on 27 and 30, respectively, of the 45 questions submitted to them. It is thus clear that, by a mere pooling of individual judgments apart from any collective operation, results would be obtained definitely superior to those of the average individual. This is in line with previous reports on pooling.

However, pooling alone can hardly account for all the collective superiority, since four of the five groups get more correct responses than such a process would give. In *C*, for example, pooling gives 52 right responses, whereas the group gets 57. Of course, it is possible that, during the short interval between taking the individual and the collective judgments, enough subjects changed their minds in the right direction to make the difference; conceivably, the time may have relieved some persons of inhibitions which blocked right associations in their first responses. Nevertheless, it seems improbable that this would occur often enough to cause any significant change.

Both observation and verbal report indicate, rather, that the increase of right responses was chiefly, if not wholly, a result of social influences in the collective situation. Doubtful subjects would often tend to delay their vocal or manual reactions just long enough to observe the dominant side, then would vote accordingly. It appeared, and it is in accord with studies of recognitive activity, that correct subjects were apt on the whole to respond somewhat more quickly and vigorously than incorrect subjects; so that in the long run the doubtful were more often carried in the right than in the wrong direction. Interesting were some of the exceptions, when a few vigorous but wrong individuals, in the midst of considerable uncertainty, would swing the vote in their direction.

It is perhaps significant that Situation 3 afforded almost no gain over the results which would have been obtained by pooling. The reason may have been the smaller size of the groups or the greater difficulty of the situation. There were three instead of two ways of voting, and splitting a group of 18 three ways leaves a possibility of only a small vote and a

small plurality, with correspondingly little social influence. Moreover, many of the items were so difficult as to create much uncertainty; in some instances a quorum would not respond, and a revote would be required. Generally the decision went in the direction of the most vigorous minority, although this might consist of only two or three persons. As a result, there were some shifts over the pluralities, invariably narrow, given by pooling, and these shifts were about as often wrong as right.

Usually not all subjects voted in the collective situation. What effect this may have had on the response it is not possible to say. Presumably, these subjects would be about equally distributed among the rights and the wrongs; although it is possible that, because of inhibitions, they were somewhat weighted on the wrong side. The matter needs further investigation.

One of the many theories of crowd behavior has it that crowds are necessarily stupid, the reason being that a crowd can not form except as its level of action drops to the level of intelligence of the most stupid participants. In other words, a collective situation requires a drop in intelligence. It is evident that such a theory does not conform to all the facts, at least such facts as are presented above. That crowd behavior is often unintelligent can not be denied. However, where this is the case, the reason would seem to lie not in the feature of collectivity as such, but rather in other elements of the situation, probably in the emotional element, which is characteristic of the common crowd.

Summary

Several aggregations of persons were presented different lists of statements on various matters of fact. Each aggregation checked the truth or falsity of the statements, operating at first individually and then as a group. The group judgments were obtained usually by a vote of acclamation. The performance of the group, in terms of frequency of correct responses, not only exceeded that of the average member but equaled or approximated that of the best member. The

existence of a probable law of collective action was indicated, namely, that, by simple collective action, the frequency of correct judgments of a group of persons will be increased approximately two sigmas over the average of their individual responses. The major reason for the group superiority was pooling of individual judgments. But analysis showed that pooling could not account for all the superiority; social influences were also operating. Uncertain members tended to be influenced to vote in the direction of the most vigorous response, and this response, in the long run, was apt to be oftener right than wrong, at least in most of the situations of this experiment.

(Manuscript received January 11, 1937)

CONFIDENCE RATINGS AS AN INDEX OF DIFFICULTY

BY MAX HERTZMAN

College of the City of New York

A number of studies have reported the tendency of confidence ratings to vary with the accuracy of performance, greater confidence being reported to go with greater accuracy. Thus, Fullerton and Cattell (**1**) reported an increase in confidence rating with an increase in the accuracy of perceiving small differences. Garrett (**2**) found a similar relationship in tasks involving the judgment of weights and the comparative length of lines. Trow's (**6**) results, based on a variety of tasks, such as recognition, immediate memory, and the judging of logical fallacies, are similar to the others.

Since tasks which are, on the average, performed accurately more frequently than others may be considered as being easier than those performed accurately with less frequency, we may from the above results conclude that relatively easy tasks would receive higher confidence ratings than relatively difficult tasks. A recognition experiment performed by Lund (**4**) tends to verify this conclusion. In the presentation period a series of nonsense syllables designated as the A group was shown to the subjects. In the test period, these syllables were mixed with two other kinds, called B and C. Syllables in the B group differed from corresponding syllables in the A group by one letter, while those in the C group differed from the A by two letters. Presumably, the more closely a syllable resembled one that was in the presentation group, the more difficult it would be to determine whether it had been in the presentation series or not. Thus, syllables of group B were more difficult to respond to than those of C. Lund found that the responses to C syllables received higher confidence ratings than those for the B.

In arranging items in an order of difficulty, the order is determined by the number of errors that the group as a whole makes on each item. Consequently, except in rare instances, regardless of its position in the difficulty scale, each item has been answered correctly by some of the individuals of the group. On the basis of a differential number of errors, the items have a differential difficulty. The question arises as to whether this difficulty value has any meaning to the individual who has answered the item correctly. Are we at liberty to say that all items that have been accurately answered by an individual are of equal difficulty to him? There may of course be some exceptional individual to whom all the items are equally easy or difficult. But it is likely that this will not be the case for the majority of subjects who will find among items which they answer correctly, various grades of difficulty. It is probable that experienced difficulty on correct items will tend to parallel difficulty measures of the same items as determined by the errors of the group. An individual may correctly answer a difficult item but he may be less certain of his accuracy than in the case of an easier item. If this is so, the concept of the relative difficulty of a particular item may be carried over even to those individuals who have answered correctly. In order to determine whether there is a significant relationship between difficulty as measured by errors and difficulty as measured by the confidence ratings of those individuals who answered correctly, the following experiment was performed.

The subjects were the students of three laboratory sections in psychology at the College of the City of New York. With two or three exceptions they were in the Sophomore or Junior year. A page bearing the photographs and names of ten different men was distributed to each subject. The subjects were given two minutes to study the pictures with the instructions that at some later interval they were to match the names and at the same time assign confidence ratings to their matchings. The three sections were treated in identical fashion with respect to all conditions with the exception of the interval between learning and matching. Group A (N = 15) was given five minutes between the two periods, group B (N = 17), fifteen minutes, and group C (N = 17), twenty-five minutes. The intervals were filled by a general class discussion of work curves. The subjects were instructed in the use of a five unit confidence scale. They were told to use 5 when absolutely certain that they were correct and 1 when equally certain that they were incorrect. The number 3 represented doubt and 4 and 2 the appropriate intermediate positions. In the test period, the subjects used a

list of names and a new sheet bearing the ten photographs without names and in a different order than in the presentation series. After each matching a confidence rating was made.

The matchings were scored and confidence ratings for correct and incorrect answers were obtained separately for each individual and for each item under each condition. As will be shown below, a transmutation technique was not only unnecessary in evaluating the confidence ratings, but may have obscured the real trends present in the data. The items were also arranged in order of difficulty as determined by the number of errors. An order of difficulty was obtained because of the combination of positional effects and the fact that even under the same conditions the names and faces were not equally easy to remember and recognize.

Since the task was similar in all three situations, in addition to obtaining results for each of the situations separately, results were also obtained using all 49 subjects as one group. This group is designated as S correlations presented below (in Table I) justify studying the combined scores as no correlation of the sub-groups with S is below .80.[1]

Rank difference correlations were obtained for the items in A, B, C, and S between the difficulty ranking and the confidence ranking where the confidence rankings were obtained only from ratings of correct matches. These correlations were .66 for A, .86 for B, .50 for C, and .83 for S. These correlations indicate that the confidence feeling of the subjects who answer correctly parallels rather closely the difficulty of the items used, and where accuracy is perfect average confidence ratings may very well become a measure of difficulty.

When the difficulty rankings of groups A, B, C, and S, and the confidence rankings of these groups are intercorrelated separately some interesting comparisons may be made. These correlations are presented in Tables I and II. The two tables though based on the intercorrelations of apparently

TABLE I

INTERCORRELATIONS OF THE DIFFICULTY VALUES

	A	B	C	S
A		.95	.68	.93
B			.55	.91
C				.80
S				

TABLE II

INTERCORRELATIONS OF THE CONFIDENCE VALUES

	A	B	C	S
A		.88	.69	.97
B			.67	.95
C				.82
S				

[1] In view also of the similarity of the situations.

different measures parallel each other to a remarkably high degree. The average absolute difference between corresponding correlations in the two tables is only .06 correlation points. This value is very small when one remembers that the median correlation in each table is over .80. Not only are the difficulty values of the items in each group closely related to the confidence ratings of the correct matchings, but these confidence ratings bear about the same relationship to each other when determined on different groups, as do the difficulty values determined on the same groups. In view of the particular conditions of this experiment, the parallel correlations would indicate that similar factors are effective in determining both the confidence values for those who answer correctly and the difficulty values of the items.

The distribution of confidence ratings has been shown to differ from individual to individual in studies such as those of Seward (**5**), Trow (**6**) and Hollingworth (**3**). This has caused a general distrust of methods which lump together confidence ratings of several individuals without first transmuting each individual's score to a common scale.[2] In the case of the present study, however, it is felt that the chief factor making for the distribution of the confidence ratings is the factor of the difficulty of the item to the individual and that this factor overshadows individual idiosyncrasies in ratings. A consideration of some additional results will show that re-scoring of the ratings was not required in this study.

1. Examination of the data shows that each individual (who does not have a perfect score) has a higher average confidence rating for his correct items than for his incorrect ones. Moreover, the average confidence rating of each of the subjects for his correct items is greater than the average confidence ratings given by the entire group for the incorrect items, taken either as a whole or individually. The average confidence rating for all the incorrect items is 2.45 which is less than the lowest individual average for correct items. When the items are taken separately the same relationship

[2] Such a scale brings with it the error of assuming that all individuals have an equal range in their confidence judgments; nevertheless the use of such a scale has been judged to be advantageous from the point of view of other considerations.

holds, as item number three, which has the largest confidence rating of all ten items, based on incorrect matchings, has an average rating of 3.22 which is still below the lowest individual average for correct items. These results would indicate that it is the individual's experienced difficulty which is largely instrumental in his use of the various parts of the scale, rather than some indefinable individual peculiarity.

2. An examination of the categories used in judging correct and incorrect items emphasizes the tendencies that are operating. With the exception of the subjects who were correct on all ten items, all the subjects contributed in some way or other to the confidence ratings of the incorrect items. Eliminating the subjects with all items correct, the comparison of the frequency with which each category was used in rating correct and incorrect answers shows some very wide differences. Incorrect items were rated *one*, 22.2 percent of the time as contrasted with 2.3 percent in the case of the correct items. For the rating of *two*, the percentages were 26.7 percent against 3.1 percent, for the rating of *three*, 38.7 percent against 7.1 percent. For the higher ratings, the tables are turned with the incorrect items receiving a rating of *four* in 11.1 percent of the cases and of *five* 1.3 percent of the times as against 17.1 percent and 69.4 percent. Clearly, then, the same people are differing widely when they are rating correct and incorrect items. For incorrect matchings, 87.6 percent of the confidence ratings fall into categories one, two, and three. For correct matchings, 86.5 percent of the ratings fall into categories four and five. Obviously, the particular extreme of the scale that is being used is not determined to any great extent by specific individual factors which are not related to the task.

3. A comparison of the confidence ratings of correct matchings for items that the errors of the group place at the extremes of the difficulty range, again points clearly to the fact that the different individuals of the group were treating the rating categories in relatively the same manner. In Table III, are presented the percentage of frequencies for each of the categories used in assigning confidence ratings for

the two easiest and the two most difficult items, *i.e.* those items which are first and second, and ninth and tenth in the difficulty scale. In addition, there is presented a similar analysis of the easy items using only those individuals who correctly matched the difficult item with which the easy one is being compared.

TABLE III

ANALYSIS OF THE DISTRIBUTION OF CONFIDENCE RATINGS FOR THE EXTREMELY EASY AND EXTREMELY DIFFICULT ITEMS (BY PERCENT)

Kind of Item	5	4	3	2	1
1. First in difficulty[3]	89.4	10.6			
2. Tenth in difficulty	34.4	28.1	18.8	9.4	9.4
3. First in difficulty using only those subjects who correctly matched the item that was tenth in difficulty	93.8	6.2			
4. Second in difficulty	89.1	6.5	2.2		2.2
5. Ninth in difficulty	45.7	25.7	14.3	8.6	5.7
6. Second in difficulty using only those subjects who correctly matched the item that was ninth in difficulty	94.3	5.7			

[3] First in difficulty is here taken to mean the easiest item. The remaining terminology should be interpreted accordingly.

One should note that all of the ratings have been obtained on items that have been matched correctly. Nevertheless, the ratings for the easier items fall predominately in the higher categories, while the percentages of such ratings for the difficult items are considerably lower. When the ratings for individuals who matched not only the easy but the difficult items correctly are studied, the same relationship holds. Individuals who were correct both in easy and in difficult items consistently, express less confidence in their matchings for the more difficult items. Again, this analysis shows that the scale used by the various individuals who were the subjects of this study could not have been fortuitous. The analysis presented in Table III corroborates the evidence of the relation of confidence ratings to difficulty, with accuracy held constant, which was furnished by the correlation coefficients.

(Manuscript received December 10, 1936)

References

1. Fullerton, G. S. and Cattell, J. M., *On the perception of small differences.* Univ. Penn. Publ., Philosophy and Science, 1892, 2, 159 p.
2. Garrett, H. E., A study of the relation of accuracy to speed, *Arch. Psych.*, 1922, No. **56**.
3. Hollingworth, H. L., Experimental studies in judgment, *Arch. Psych.*, 1913, No. **29**.
4. Lund, F. H., The criteria of confidence, *Amer. J. Psych.*, 1926, **37**, 372–381.
5. Seward, G. H., Recognition time as a measure of confidence, *Arch. Psych.*, 1928, No. **99**.
6. Trow, W. C., The psychology of confidence, *Arch. Psych.*, 1923, No. **67**.

REPORT OF A MINOR INVESTIGATION OF EXTRA-SENSORY PERCEPTION

BY KENNETH H. BAKER

University of Minnesota

In view of the fact that there has been collected a considerable amount of evidence which appears to demonstrate the existence of what Dr. J. B. Rhine has chosen to call Extra-sensory Perception (E.S.P.) it seems only natural that this phenomenon should be the subject of a good deal of laboratory experimentation. That this problem has received serious attention in several laboratories throughout the country is well known, although the results of only two or three of these investigations have come to the writer's attention.[1]

The problem of extra-sensory perception was presented to a large class of freshmen and sophomores at the University of Minnesota. As a result of a discussion of the problem before the class in which Rhine's results were briefly reviewed, it was believed that a favorable, coöperative attitude was established in a majority of the students of the class. At this point in the discussion, the class was asked to participate in an attempt to determine whether or not there were people in the group who could approximate or duplicate Rhine's results under the conditions of the classroom. Slips of paper were distributed, one to each student. Each student was asked to put his name at the top of the slip and a mark (such as an X) near his name if he expected to make a better-than-chance score in the test. This mark was supposed to have been made before the test itself was begun. We have

[1] It is the writer's conviction that the results of all experiments on extra-sensory perception, positive, equivocal and negative, should be published so that the literature will not become biased. For this reason, the results presented in this report are being published even though they are equivocal and represent nothing more than a preliminary attempt to secure subjects who could duplicate the performance of Rhine's subjects with the Zener cards.

no reason to suspect that any of these marks were changed after the test. The deck of Zener cards was held by the instructor who stood on a platform in the front of the room in full view of each student. The method used was that which is designated by Rhine as PC BT-25,[2] that is, the deck was held face down and the cards removed from the top of the deck one at a time after the class had been given an opportunity to call each card and record the response. After the entire deck had been called once, it was gone through again by the instructor who called out the name of each card so that the students could check their responses and record the number of 'hits.' The frequency distribution of scores obtained in the class test is shown in Table I together

TABLE I

FREQUENCY OF 'HITS' IN CLASS EXPERIMENT

Hits	f_o*	f_t†	X‡
12	0	0	
11	0	1	
10	3	3	
9	6	7	1
8	16	15	
7	28	26	1
6	39	39	3
5	36	46	1
4	38	44	1
3	34	33	5
2	20	17	
1	15	6	1
0	2	1	

Mean = 4.77.

* Obtained frequency.

† Theoretical frequency.

‡ Expected to receive high score before test.

with a theoretical frequency for each number of 'hits' and a distribution of the scores made by those "who expected to make a better-than-chance score."

After the scoring had been completed, the instructor asked for a show of hands of those who had made eight or more hits. When it was apparent that this number was more than twenty, these students were urged to come to the laboratory, at a time convenient to them, to participate in further tests.

[2] J. B. Rhine, *Extra-Sensory Perception*, 1935, Boston: Bruce Humphries, p. xiv.

The assumption that a majority of the class was coöperative in this test is somewhat justified by the fact that eighteen of the twenty-five people who scored eight or better subsequently came to the laboratory for additional tests. The data from this testing procedure are tabulated in Table II.

TABLE II

RECORD IN FURTHER TESTS OF THOSE RECEIVING SCORE OF 8 OR MORE IN CLASS EXPERIMENT

Subj.	Hits in Class	No. Called	No. Hits	Necess. for Signif.	Av. per 25	Meth. Used	Trials Witnessed
WD	10	250	44	66	4.4	BT-25	Yes
RE	10	250	59	66	5.9	"	"
TTW	9	250	52	66	5.2	"	"
RJW	9	250	60	66	6.0	"	"
MN	9	250	53	66	5.3	"	"
RJN	9	250	52	66	5.2	"	"
BJ	9	250	60	66	6.0	"	"
RPD	9	250	57	66	5.7	"	"
MMM	8	250	45	66	4.5	"	"
JPK	8	250	49	66	4.9	"	"
CC	8	200	34	55	4.25	"	"
DV	8	250	56	66	5.6	"	"
WBH	8	250	45	66	4.5	"	"
FCG	8	250	39	66	3.9	"	"
RWC	8	200	42	55	5.25	"	"
CM	8	250	57	66	5.7	"	"
KR	8	250	50	66	5.0	"	"
RC	7	250	47	66	4.7	"	"
Totals		4400	901	960	5.1		

It is unfortunate that the three people who made the highest scores in the later tests (RE, RJW, BJ) could not be tested further. These subjects did not believe that they possessed any special ability and attributed their scores to 'just luck.' Lack of time on the part of the students was another cause for failure to keep later appointments. The tests were conducted during the last weeks of the spring quarter.

In further attempts to uncover suitable subjects, the writer has announced in all of his classes that he is looking for such subjects and that volunteers would be most welcome. Ten people have responded in an effort to determine whether or not they could demonstrate extra-sensory ability. The data from these subjects are recorded in Table III.

TABLE III

PERFORMANCE OF A HETEROGENEOUS GROUP OF VOLUNTEER SUBJECTS

Subj.	No. Called	No. Hits	Necess. for Signif.	Av. per 25	Meth. Used	Trials Witnessed
AJ.....	125	24	37	4.8	BT-25	No
AJ.....	200	43	55	5.6	"	Yes
ER.....	175	91	49	13.0	"	No
HV.....	125	25	37	5.0	"	No
HV.....	200	38	55	4.75	BT-5	Yes
HV.....	25	3	11	3.0	BT-1	No
RH.....	100	21	30	5.25	BT-25	Yes
CB.....	175	41	49	6.83	"	Yes
JR.....	250	47	66	4.7	DT	No
JR.....	100	15	30	3.75	BT-5	Yes
JR.....	100	14	30	3.5	DT	Yes
GHC.....	175	35	49	3.89	BT-5	Yes
JH.....	250	43	66	4.3	BT-25	No
JRB.....	500	79	123	3.95	BT-5	Yes
WC.....	75	16	24	5.33	BT-5	Yes
WC.....	50	12	18	6.0	BT-25	Yes
Totals.....	2625	547	585	5.20		

One of the writer's students became interested in the problem and volunteered to collect data among her friends and acquaintances, many of whom at some time or other had laid claim to a peculiar 'psychic' ability. The writer believed that this student would be better suited than himself to collect the data because of the greater opportunity for full coöperation on the part of the subjects. The student was thoroughly acquainted with the problem, the controls necessary and the procedures that should be used. The data from these subjects are recorded in Table IV.

Another procedure suggested by Rhine in this connection has also been used. A number (from 0 to 9) was written in black ink on a piece of black cardboard and the cardboard placed in a plain envelope. The envelope was sealed and stapled and identified on the outside by a letter. The actual numbers were not known to the writer since the identifying letters were put on the envelopes after they had been sealed and thoroughly shuffled. More than fifty students were given the opportunity of trying to 'guess' what number was written on the card inside these envelopes. None of these

TABLE IV

PERFORMANCE OF A HETEROGENEOUS GROUP OF VOLUNTEER SUBJECTS (DATA COLLECTED BY STUDENT D.A.R.)

Subj.	No. Called	No. Hits	Necess. for Signif.	Av. per 25	Meth. Used	Trials Witnessed
DM.....	50	9	18	4.5	BT-25	Yes
DM.....	25	8	11	8.0	DT	"
MM.....	100	22	30	5.5	BT-25	"
MM.....	125	28	37	5.6	DT	"
MM.....	50	8	18	4.0	BT-5	"
AWR.....	75	8	24	2.67	BT-25	"
AWR.....	100	17	30	4.25	DT	"
CP.....	175	42	49	6.00	BT-25	"
CP.....	50	8	18	4.00	BT-5	"
CP.....	50	6	18	3.0	DT	"
ICP.....	75	17	24	5.67	BT-25	"
ICP.....	100	22	30	5.5	DT	"
SSC.....	175	38	49	5.43	BT-25	"
MAP.....	200	40	55	5.00	"	"
MAP.....	50	9	18	4.5	BT-5	"
MAP.....	50	10	18	5.0	DT	"
BK.....	125	21	37	4.2	BT-25	"
BK.....	25	5	11	5.0	DT	"
MB.....	125	16	37	3.2	BT-25	"
MB.....	50	14	18	7.0	DT	"
BL.....	100	19	30	4.75	BT-25	"
BL.....	50	13	18	6.5	BT-5	"
BL.....	125	26	37	5.2	DT	"
Total.....	2050	406	464	4.95		

students received scores above those that may be accounted for by chance.

As was stated earlier in this report, our primary purpose in conducting the tests described herein was an attempt to locate among the students on this campus some who could duplicate or better the performance of Rhine's subjects at Duke University. Our failure to do so is quite definitely indicated by the data. The testing conditions of the class experiment were far from ideal, although they do fulfill one of the qualifications stated by Rhine, namely, that when others are in the same room with the subject they should be occupied with the same task or with another task that will not disturb the subject. Throughout the tests, the writer felt that the subjects were coöperating to their fullest extent. Although some of the subjects 'did not believe in' the existence of the ability under investigation or claimed that they did

not possess it, there seemed always to be the attempt to make as high a score as possible. The laboratory conditions for the additional tests seemed, on the other hand, to be close to ideal. The subjects appeared for testing whenever they chose, the room was quiet, the test was carried on without interruption, various individual motivating devices were used, and the test was stopped whenever the subject requested, although each subject was encouraged to run through the deck at least ten times.

We may conclude that none of the subjects tested displayed an extra-sensory ability.

(Manuscript received January 9, 1937)